



Université de Toulouse
Institut National Polytechnique de Toulouse
École Nationale Supérieure Agronomique de Toulouse

RAPPORT DE PROJET DE FIN D'ÉTUDES

Pour l'obtention du

Diplôme d'Ingénieur Agronome
Spécialité : Agrogéomatique

**Pratiques agricoles, paysage et
régulation naturelle des insectes
ravageurs des cultures :
Etude de cas dans les Niayes au Sénégal**

GUERIN Solange



Maître de stage : Mme Valérie SOTI et M. Thierry BREVAULT
Tuteur-enseignant : M. David SHEEREN

Septembre 2013

Remerciements

- Mme Valérie Soti, mon maître de stage, pour l'aide précieuse qu'elle a su m'apporter dans le domaine de la télédétection et de l'analyse statistique.
- Mr Thierry Brévault, mon maître de stage, pour l'aide non moins précieuse dans le domaine de l'entomologie.
- L'équipe de l'UMR TETIS et la Maison de la Télédétection pour m'avoir soutenu et supporté pendant la partie française de ce stage.
- Chacun des membres de l'équipe de travail sénégalaise pour m'avoir accompagné et prêté main forte lors des phases de terrain au Sénégal.
- Dr Karamoko Diarra, Professeur à l'UCAD, pour les moments riches en enseignements qu'ont été nos discussions.
- L'équipe Carabe en particulier et tous les chercheurs du CIRAD en général, qui ont émaillé mon stage de très belles rencontres.
- Mr Sheeren, mon tuteur ENSAT, pour son soutien à toutes les étapes de ce cheminement, pour son aide et ses conseils avisés
- les équipes pédagogiques de l'ENSAT et de l'Université Toulouse Mirail pour nous avoir menés au terme de cette année de formation de manière si efficace et de nous avoir permis d'acquérir un bagage vraiment utile.

Résumé

Le contrôle des ravageurs est une composante importante de la conduite des cultures permettant d'assurer le rendement et la qualité de la production. La dynamique des populations de ravageurs (ici des insectes) dépend de nombreux paramètres, dont les pratiques agricoles à la parcelle (traitements phytosanitaires), la régulation biologique (via les ennemis naturels), mais aussi le contexte paysager (composition et disposition de l'espace). Cette étude cherche à établir l'effet des pratiques culturales et du contexte paysager sur l'abondance de la teigne du chou *Plutella xylostella*, et la régulation par les ennemis naturels (parasitoïdes) dans la zone des Niayes au Sénégal. Pour mener à bien cette étude, il a été nécessaire de sélectionner 31 parcelles d'observation, de collecter des données sur les pratiques culturales et sur les insectes ravageurs, puis d'établir une classification de l'image satellite de la zone étudiée afin d'en extraire les variables environnementales. Enfin, une analyse statistique a permis de déterminer les variables qui expliquent le mieux le pourcentage de choux infestés et le taux de parasitisme des insectes collectés. Cette étude montre que le taux d'infestation augmente avec la complexité du paysage et avec le nombre de traitements insecticides tandis que le contrôle biologique par les parasitoïdes augmente avec la présence de plantes-hôtes et de vergers dans le paysage.

Abstract

Pest control is an important component of crop management to insure yield and quality of production. Dynamics of the pest populations (here insects) depends on various parameters, such as cultural practices (insecticide spraying), biological control (via natural enemies), but also landscape context (composition and configuration of landscape). This study aims at establishing the effect of cultural practices and landscape context on the abundance of the diamondback moth (*Plutella xylostella*) and control by natural enemies (parasites) in the Niayes area in Senegal. To successfully carry out this study, it was necessary to select 31 focal fields, to collect data on cultural practices and target insects, then to establish a classification of the satellite images of the studied area to extract landscaped variables. Finally, a statistical analysis was conducted to determine variables explaining the percentage of infested cabbages and the rate of parasitism. This study showed that cabbage infestation is related to landscape context and to the insecticide use while biological control by parasitoids depends on landscape composition and selectivity of insecticides used.

Key words : Integrated Pest Management, biological control, landscape, agricultural practices, Niayes, Senegal, diamondback moth, *Plutella xylostella*

Sommaire

REMERCIEMENTS	2
RESUME	3
ABSTRACT	3
INTRODUCTION	7
I. CADRE DU PROJET	8
A. LE CIRAD	8
1. Missions du CIRAD	8
2. Organigramme du CIRAD	9
3. L'UPR SCA	10
B. PROJET BIOBIO	10
II. ETAT DE L'ART ET PROBLEMATIQUE	11
A. ETAT DE L'ART	11
1. Connaissances biologiques	11
2. Connaissances sur le paysage	13
3. Connaissances sur les pratiques agricoles	13
4. Synthèse des connaissances	14
B. PROBLEMATIQUES	14
C. DEROULEMENT DU STAGE	15
D. DESCRIPTION DE LA ZONE D'ETUDE	16
1. Les Niayes en général	16
2. La zone d'étude en particulier	16
III. MATERIELS	17
A. CARACTERISTIQUES DES IMAGES SATELLITES PLEIADE	17
1. Présentation des images satellites Pléiades	17
2. Caractéristiques des images Pléiades utilisées	17
B. LOGICIELS	18
1. Logiciels de traitements de l'image	18
2. Logiciels d'acquisition, de gestion et d'analyse des données	19
IV. DEMARCHE METHODOLOGIQUE	19
A. PHASE DE TERRAIN ET LABORATOIRE	20
1. Choix des parcelles	20
2. Capture des insectes	21
3. Suivi en laboratoire	21
B. TRAITEMENT DE L'IMAGE	22
1. Prétraitements des images satellites	22
2. Segmentation et classification de l'image	23
3. Différents niveaux de classification	26
4. Collecte des données environnementales	27
5. Validation de la classification	28
6. Choix des variables environnementales	29
C. ANALYSE DES DONNEES	30
1. Compilation sous Access	30
2. Analyse statistique sous R	32
V. ANALYSE DES RESULTATS	35
A. RESULTATS ISSUS DES DONNEES DE TERRAIN ET DE LABORATOIRE	35
1. Parcelles	35
2. Individus	37
3. Pratiques agricoles	39
B. RESULTATS DU TRAITEMENT DE L'IMAGE SATELLITE	40
1. Matrice de confusion	40
2. Causes d'erreur de la classification	41
3. Classification finale	42
C. EFFET DES PRATIQUES CULTURALES ET DU PAYSAGE SUR LE TAUX D'INFESTATION ET LE TAUX DE PARASITISME	43
1. Connaissances des variables	43

2. Etude des corrélations entre les variables	43
3. Recherche des modèles selon la loi normale.....	45
4. Recherche des modèles selon la loi binomiale.....	47
DISCUSSION	52
CONCLUSION.....	54

Liste des annexes

Annexe 1 : Table récapitulative des données de sélectivité.....	57
Annexe 2 : Schéma de transformation de la valeur des pixels du compte numérique vers la valeur de luminance	58
Annexe 3 : Schéma de la transformation de la valeur des pixels de la valeur de luminance vers la valeur de réflectance	58
Annexe 4 : Paramètres de tests de la segmentation.....	60
Annexe 5 : Arbre de classification, arbre dichotomique de décision et table des paramètres de classification.....	61
Annexe 6 : Glossaire des variables utilisées pour l'étude statistique.....	64
Annexe 7 : Variables biologiques utilisées dans l'étude statistique.....	65
Annexe 8 : Variables environnementales utilisées pour l'analyse statistique.....	66
Annexe 9 : Variables relatives aux pratiques culturelles utilisées pour l'analyse statistique ...	67
Annexe 10 : Image globale après classification selon le niveau 2 de regroupement	68
Annexe 11 : Histogramme et données des variables de l'étude statistique	69
Annexe 12 : Table des corrélations entre les variables V1 et V3 et les variables explicatives	72
Annexe 13 : Table de corrélation des variables explicatives	73

Liste des Tableaux

Tableau 1 : Caractéristiques physiques des images Pléiades	18
Tableau 2 : Caractéristiques spectrométriques des images Pléiades.....	18
Tableau 3 : Fiche de suivi des individus de <i>Plutella xylostella</i> en laboratoire	22
Tableau 4 : Description des différents niveaux de classification	26
Tableau 5 : Nombre de points de référence alloués par classe.....	28
Tableau 6 : Extrait de la table attributaire après jointure	29
Tableau 7 : Matrice de distribution des points	29
Tableau 8 : Matrice de confusion pour la classification de niveau 2	41
Tableau 9 : Indices de classification.....	41
Tableau 10 : Table des corrélations et p-value des variables à expliquer	43
Tableau 11 : Modèles des pratiques culturelles pour V1	45
Tableau 12 : Modèles des variables du paysage pour V1	45
Tableau 13 : Modèles des pratiques culturelles pour V3 (linéaire).....	46
Tableau 14 : Modèles des variables du paysage pour V3 (linéaire).....	46
Tableau 15 : Modèles de pratiques culturelles pour V1 (binomial)	47
Tableau 16 : Modèles de variables de paysages pour V1 (binomial).....	47
Tableau 17 : Poids des variables dans le modèle binomial de V1	48
Tableau 18 : Modèles de pratiques culturelles pour V3 (binomial).....	49
Tableau 19 : Modèles des indicateurs paysagers pour V3 (binomial)	50
Tableau 20 : Poids des variables dans le modèle binomial de V3	50

Liste des figures

Figure 1 : Le CIRAD dans le monde.....	8
Figure 2 : Organigramme du CIRAD.....	9
Figure 3 : Classification et photo de <i>Plutella xylostella</i>	11
Figure 4: Cycle de vie de <i>Plutella xylostella</i>	12
Figure 5 : Classification et photo de chou pommé.....	12
Figure 6 : Principaux parasitoïdes de <i>Plutella xylostella</i> , A: <i>Oomyzus skolowskii</i> , B: <i>Brachymeria sp.</i> , C: <i>Apanteles litae</i> et D: <i>Cotesia plutellae</i>	12
Figure 7 : Schéma récapitulatif des connaissances	14
Figure 8 : Diagramme de Gantt du stage.....	15
Figure 9 : Localisation de la zone des Niayes	16
Figure 10 : Localisation de la zone d'étude (en vert, source : Word Street Map)	16
Figure 11 : Logo ERDAS	18
Figure 12 : Logo E cognition	18
Figure 13 : Logo ArcGIS	19
Figure 14 : Logo ArcPad	19
Figure 15 : Logo Access.....	19
Figure 16 : Logo R	19
Figure 17 : Barème d'attribution de la note d'adventice en fonction du pourcentage de surface recouverte par des adventices.....	20
Figure 18 : Exemple de protocole de choix des plants à échantillonner (sur 4 bandes)	21
Figure 19 : Zones utilisées pour la création de l'image test (A) et image test obtenue (B).....	23
Figure 20 : Boîte de dialogue du logiciel e-cognition pour le choix des paramètres de segmentation.....	24
Figure 21 : Résultats de la segmentation (A : image brute et B : image segmentée).....	25
Figure 22 : Arbre de classification	25
Figure 23 : Exemple de classification (A : image brute, et B : image classifiée)	26
Figure 24 : points de références utilisés pour la validation de la classification	28
Figure 25 : Formule de l'indice de Shannon.....	30
Figure 26 : Relations entre les tables de la base de données Access.....	31
Figure 27 : Code R pour l'appel des données et des librairies.....	32
Figure 28 : Code R pour l'étude des corrélations entre les variables	33
Figure 29 : Code R pour la recherche des modèles selon la loi gaussienne.....	34
Figure 30: Codes R pour la recherche de modèles selon la loi binomiale	34
Figure 31 : Répartition spatiale des parcelles observées.....	35
Figure 32 : Evolution de la taille des choux selon le numéro d'observation.....	36
Figure 33 : Répartition du nombre moyen de larves par plant.....	36
Figure 34 : Répartition du taux d'infestation par parcelle	36
Figure 35 : Répartition spatiale de l'infestation.....	37
Figure 36 : Répartition des espèces observées	37
Figure 37 : Répartition des individus observés selon le stade de collecte	38
Figure 38 : Infestation parasitaire en fonction du stade du ravageur à la collecte	38
Figure 39 : Type de parasites impliqués dans l'infestation.....	38
Figure 40 : Précédent cultural des différentes parcelles.....	39
Figure 41 : Type de sol des différentes parcelles	39
Figure 42 : Nombre de traitements phytosanitaires réalisés après le repiquage par parcelle...	40
Figure 43 : Type de produits phytosanitaires utilisés.....	40
Figure 44 : Classification de l'image totale	42
Figure 45 : étude graphique du modèle prédictif pour V1	49
Figure 46: Etude graphique du modèle prédictif pour V3	51

Introduction

Le projet de fin d'étude est, pour un étudiant, la cristallisation des enseignements acquis pendant sa scolarité et le dernier pas avant le monde du travail. C'est pourquoi le choix de la structure d'accueil de ce projet est important dans la suite du projet de vie et de carrière envisagé à la sortie de l'école. Même si ce projet n'est pas encore totalement focalisé autour d'un métier pour moi, il me semble évident que mon envie de carrière se porte vers la recherche et le développement de projets dans un but d'aide aux autres (gestion de crise, aide au développement...).

Le projet qui m'a été proposé par le CIRAD est en relation avec ma formation initiale et ma spécialisation de géomatique puisqu'il concerne l'étude du lien entre les pratiques agricoles, l'environnement et la régulation des insectes ravageurs des cultures de choux à travers l'utilisation des images satellites couplées à des enquêtes sur le terrain (zone des Niayes, Sénégal). Ce stage est centré sur l'étude de l'effet des pratiques agricoles et du contexte paysager sur les infestations de cultures de choux par les insectes ravageurs. L'objectif du stage est donc de caractériser précisément les paramètres régulant la présence des ravageurs, en vue de développer des modèles prédictifs et d'améliorer la conduite des cultures et l'aménagement du territoire pour optimiser la productivité de la parcelle et le respect de son environnement.

En Afrique sub-saharienne, les cultures maraîchères constituent une composante essentielle du développement durable, de par une contribution importante à la sécurité et à l'équilibre alimentaire, mais aussi une source importante de revenus et d'emplois pour les populations les plus pauvres. Cependant, les bio-agresseurs, en particulier les insectes ravageurs, sont un frein majeur à l'amélioration de la productivité et de la qualité de la production horticole. La lutte chimique est la principale stratégie de lutte mise en œuvre par les maraîchers pour protéger leurs cultures, avec de fréquentes dérives dans l'utilisation des insecticides, en particulier lorsque ceux-ci perdent leur efficacité suite au développement de résistance chez les ravageurs ciblés. Ces « mauvaises » pratiques affectent la sécurité sanitaire des produits (résidus de pesticides) mais aussi la santé des producteurs et de l'environnement (biodiversité, ennemis naturels, services éco systémiques). La culture du chou est, par exemple, devenue impossible dans certaines zones des Niayes (Pikine) à force d'attaques incontrôlables de la teigne du chou, *Plutella xylostella*. C'est pour cela que l'étude porte sur ce ravageur particulier.

Pour présenter le travail réalisé durant le stage, il est d'abord nécessaire de faire un point sur l'organisme d'accueil et les équipes qui le composent. Ensuite, il est important d'explicitier la problématique et les objectifs de ce stage. Conformément au plan de tout article de recherche, il est justifié de développer le matériel puis la méthode mis en œuvre dans le projet avant de détailler les résultats obtenus. Pour finir, une discussion sur le projet et ses résultats devra être menée afin de conclure sur la question développée durant le stage.

I. Cadre du projet

Le cadre d'un projet intègre autant l'organisme d'accueil du projet que la zone sur laquelle il s'applique. Le lien entre les deux permet de mieux apprécier la légitimité et l'intérêt de l'étude menée. Il est donc important de s'étendre sur ces deux points afin de cerner au mieux les contraintes liées à l'environnement qu'il soit physique ou institutionnel.

A. Le CIRAD

L'Organisme d'accueil de ce stage est le CIRAD. Le Centre de Coopération International de Recherche Agronomique pour le Développement est un organisme de recherche créé sous la forme d'un établissement public à caractère industriel et commercial (EPIC) en 1984¹. Neuf instituts de recherche agricole tropicale sont à l'origine de cet organisme, regroupant ainsi leur spécialité dans un type de production particulier (coton, oléagineux, maraichage, caoutchouc...) en une structure unique et plus polyvalente. Cette association a permis de mutualiser les fonds, les moyens et les connaissances au service de la recherche dans les pays du Sud.

L'ensemble des informations décrites dans cette partie sont extraits du site internet du CIRAD et des sites des différentes unités de recherches. Grâce à ces informations, il est possible de bien cerner les missions que cet organisme défend, ainsi que le cadre opérationnel du stage afin de définir et de comprendre le lien précis entre le sujet du stage et l'entreprise qui l'encadre.

1. Missions du CIRAD

Le CIRAD concentre ses recherches autour de 6 axes scientifiques prioritaires² regroupant des thématiques autour de l'alimentation humaine, de l'intensification agricole au service de l'environnement et des sociétés, de la santé animale et de la lutte contre la pauvreté et les inégalités.

Le sujet de ce stage s'intègre dans l'axe scientifique concernant l'intensification écologique puisque le cœur de cet axe repose sur la conception de systèmes à la fois plus productifs, plus respectueux de l'environnement et moins gourmands en intrants. Cet axe mène également à tenter d'inventer de nouvelles techniques de lutte contre les maladies et les ravageurs et comprendre le fonctionnement de la nature afin d'en exploiter les ressources sans la détruire.

Comme le montre la figure 1, le CIRAD³ dispose d'une représentation importante sur les territoires sahéliens et tropicaux de la planète ainsi que d'un réseau mondial de partenaires afin d'agir au plus près des populations.

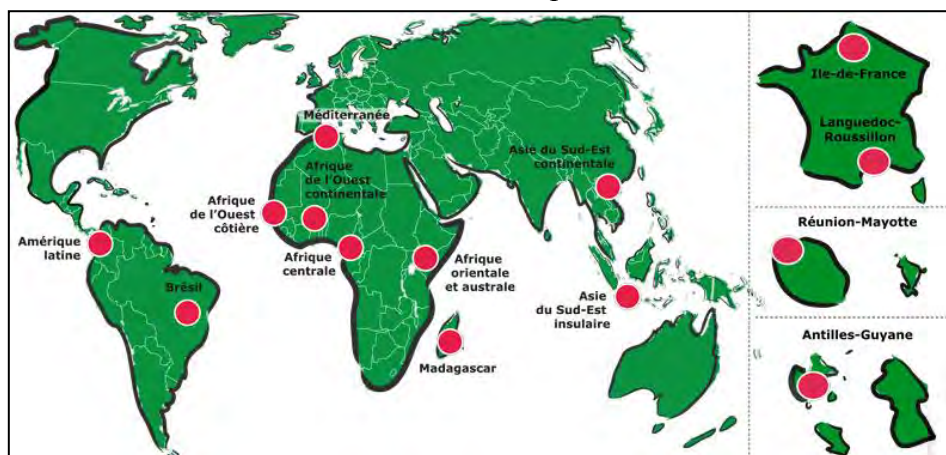


Figure 1 : Le CIRAD dans le monde

¹ Site internet du CIRAD, rubrique 'en bref'

² Site internet du CIRAD, rubrique 'axes prioritaires'

³ Site internet du CIRAD rubrique 'Le CIRAD dans le Monde'

Ainsi, des activités de coopération sont menées dans plus de 90 pays, en plus de la mise en place d'un centre de recherche et de formation à Montpellier en appui aux délégations régionales. C'est grâce à cette structuration en direction régionale, qu'il m'a été possible d'effectuer les recherches de terrain au Sénégal.

Après les missions et la répartition géographique des agents du CIRAD, il est nécessaire de s'intéresser à la structure interne d'un tel établissement, afin d'en comprendre au mieux les rouages et les mécanismes de fonctionnement.

2. Organigramme du CIRAD

Le CIRAD compte 1800 employés, dont 800 chercheurs répartis sur le globe.

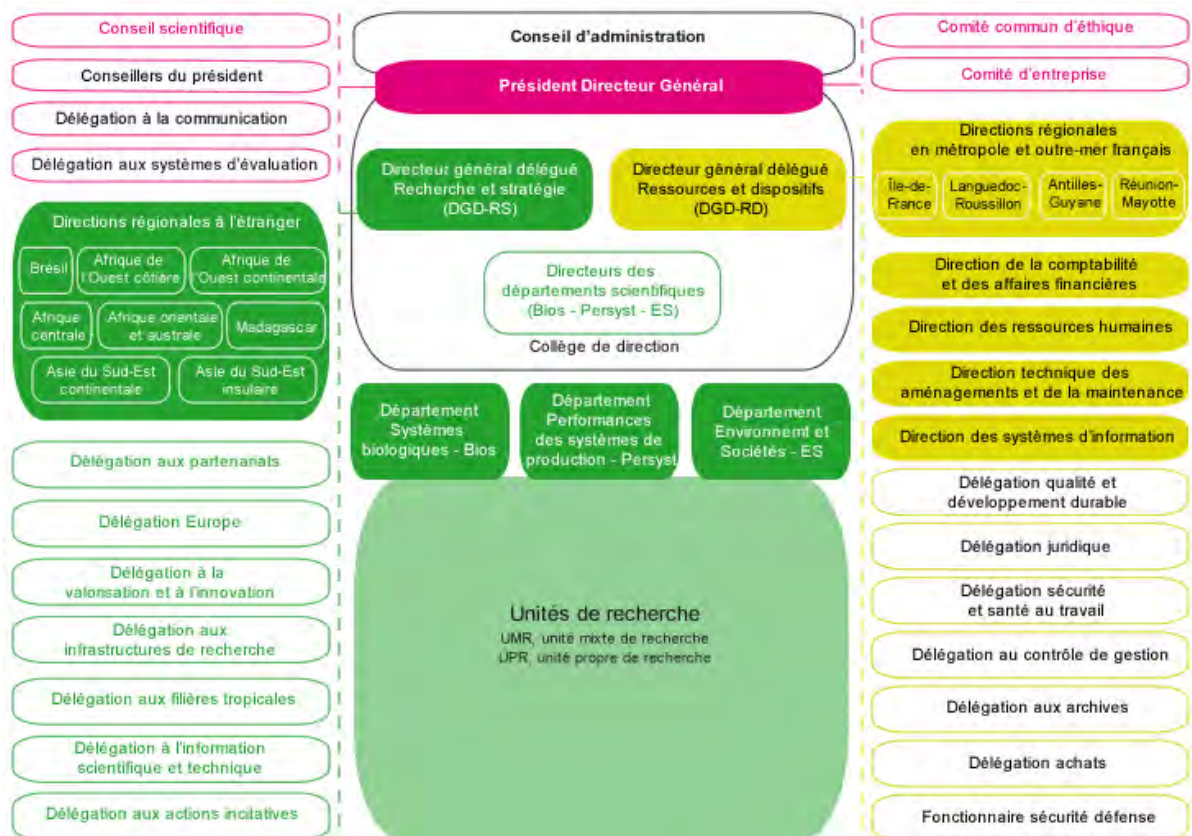


Figure 2 : Organigramme du CIRAD

Comme le présente la figure 2, l'organigramme du CIRAD peut être divisé en trois grands groupes représentant chacun un pôle d'activités. Le premier est matérialisé par les bulles de couleur rose et représente le pôle purement décisionnel du CIRAD avec la présidence et les instances et comités de conseil et de décision. Le deuxième pôle est celui des fonctions administratives (en jaune) assurant le fonctionnement correct des différents services sur les plans financiers, informatiques ou encore humains. Pour finir, les bulles en vert regroupent l'ensemble des activités de recherche, dont les différents départements de recherche et les unités de recherche proprement dites.

Une fois la structure générale de l'organisme étudiée, il est important de savoir avec précision dans quelle cadre se déroule le stage, afin de bien établir le lien entre le sujet et les thèmes de recherche des équipes encadrantes.

3. L'UPR SCA

Le projet proposé comme sujet pour ce stage est inscrit dans le cadre de la recherche menée par l'UPR SCA (Système de Culture Annuelle). Cette unité propre de recherche est sous la tutelle du département Persyst (Performances des systèmes de production et de transformation tropicaux), et étudie les processus biologiques en interaction avec l'environnement et les pratiques techniques.

L'unité SCA s'intéresse de près à l'agroécologie et à l'intensification durable des cultures annuelles. En effet, cette unité se penche sur la conception de système de production à la fois performant et durable pour les cultures annuelles (céréales, maraichages...) pour permettre aux paysans des pays en voie de développement de valoriser leur production. Cette unité propre de recherche regroupe cinq équipes, dont l'équipe Carabe (Caractérisation et gestion intégrée des risques d'origine biotique) où j'ai effectué mon projet de fin d'étude. L'équipe Carabe s'intéresse à la caractérisation et à la gestion intégrée des risques d'origine biotique. Durant ce stage, j'ai été hébergée par l'UMR TETIS (Territoires, Environnement, Télédétection et Information Spatiale) à la Maison de la Télédétection de Montpellier dotée d'infrastructures et de matériels informatiques performant en matière de traitement et d'analyse de l'informations environnementales dérivées des images satellites..

B. Projet BioBio

Le projet BioBio (Biodiversité et gestion des bio-agresseurs dans les paysages agricoles) est un projet soutenu financièrement par l'AIRD (Agence Inter-établissement de Recherche pour le Développement). Ce financement s'inscrit dans le cadre du programme PEERS (Programme d'Excellence pour l'Enseignement et la Recherche au Sud) afin de renforcer le partenariat entre deux chercheurs confirmés, un chercheur du Nord et un enseignant chercheur du Sud.

Les partenaires de ce projet sont le CIRAD ainsi que deux structures sénégalaises, l'ISRA-CDH (Institut Sénégalais de Recherche Agricole - Centre de Développement de l'Horticulture) et l'UCAD (Université Cheik Anta Diop de Dakar) avec le Master Gedah (Gestion durable des agro-écosystèmes horticoles).

Ce partenariat contribue au renforcement de la communauté scientifique du Sud et à l'animation de Dispositif Prioritaire DivEcoSys (Diversification des systèmes de culture et gestion agro-écologique des bio-agresseurs en Afrique de l'Ouest⁴) porté par le CIRAD.

Une fois le cadre professionnel posé, il est nécessaire de décrire le contexte et les objectifs du stage, afin de mieux appréhender les enjeux liés à la problématique de la régulation des insectes ravageurs des cultures du chou au Sénégal.

⁴ <http://divecosys.e-monsite.com/>

II. Etat de l'art et problématique

La première étape dans la compréhension des enjeux et de la problématique qui motive la mise en place d'une étude est l'analyse bibliographique du sujet. La seconde étape est la définition précise de la problématique et des objectifs du stage.

A. Etat de l'art

La connaissance de l'état de l'art est importante afin de savoir quel sont les connaissances sur lesquelles s'appuie l'étude. Ainsi, certaines connaissances préalablement acquises permettront de poser des hypothèses de travail cohérentes vis à vis des connaissances déjà existantes dans le domaine. L'ensemble des connaissances relatives au sujet du stage peut être divisé selon les trois termes majeurs de ce sujet, 'insectes ravageurs', 'environnement' et 'pratiques culturelles'. C'est donc en suivant ce plan que seront détaillées les connaissances concernant la biologie du ravageur ciblé, des ravageurs et les parasitoïdes qui y sont associées, les connaissances concernant le paysage et celle concernant les pratiques agricoles. Il sera enfin fait une synthèse de ses trois jeux de connaissances.

1. Connaissances biologiques

La biologie des insectes ravageurs des cultures ne peut être pleinement appréhendée sans connaître les autres composants de ce cycle de vie. En effet, ces insectes se nourrissent de plantes hôtes appartenant à des espèces précises (ici les crucifères) et sont eux-mêmes menacés par des parasitoïdes. C'est donc selon le schéma ravageur-hôte-parasitoïde que sera développée cette partie.

1). Les insectes ravageurs

Même si plusieurs insectes sont responsables de dégâts dans les cultures de chou au Sénégal, l'insecte ciblé pour cette étude est *Plutella xylostella*. Appartenant à l'ordre des lépidoptères (voir figure 3), il cause des dégâts au stade larvaire en se nourrissant des feuilles (Furlong et al., 2008).

Les larves commencent leur développement en dévorant les feuilles. Lors de la nymphose, les larves produisent des chrysalides d'environ 8 mm de couleur blanchâtre. Au stade adulte, le papillon mesure entre 12 et 17 mm. (Talekar and Shelton, 1993) et est capable d'effectuer des migrations à grande distance. Lors des périodes sèches, trois à six générations peuvent se succéder (voir figure 4).

Figure 3 : Classification et photo de *Plutella xylostella*

Nom commun : teigne du chou
Nom scientifique : *Plutella xylostella*

Règne : Animalia
Embranchement : Arthropoda
Sous-embranchement : Hexapoda
Classe : Insecta
Sous-classe : Pterygota
Super ordre : Neoptera
Ordre : Lepidoptera
Sous-ordre : Microlepidoptera
Super famille : Yponomeutoidea
Famille : Plutellidae
Genre : *Plutella*



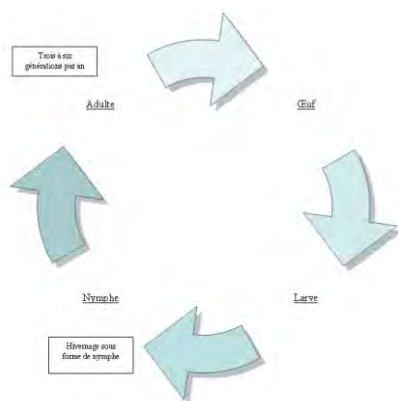


Figure 4: Cycle de vie de *Plutella xylostella*

parasitoïdes. Ici, seuls les parasitoïdes s'attaquant à l'espèce *Plutella xylostella* seront présentés.

2). Les plantes hôtes

Nom commun : Chou pommé
Nom scientifique : *Brassica capitata*

Règne : Plantae
Sous règne : Tracheobionta
Division : Magnoliophyta
Classe : Magnoliophyta
Sous-classe : Dilleniidae
Ordre : Brassicales
Famille : Brassicaceae
Genre : *Brassica*



Figure 5 : Classification et photo de chou pommé
développement du chou.

Dans le cadre de l'étude, deux plantes hôtes sont considérées, le chou commun (voir figure 5) et le navet (*Brassica rapa*). Même s'il est intéressant comme hôte secondaire potentiel, aucun prélèvement d'insectes n'a été réalisé sur le navet. Le chou est ici la plante référence sur laquelle est effectué l'ensemble des mesures et prélèvements.

Le chou est cultivé pour ses feuilles qui se regroupent en pomme à la fin de la phase végétative. Théoriquement vivace (durée de vie de plus d'un an), le chou est principalement cultivé comme une plante annuelle et récolté avant la floraison (Azanza et al., 1997). La croissance totale du chou entre le stage du semis et celui de la récolte n'excède que rarement les 80 jours.

Les attaques d'insectes ravageurs se déroulent essentiellement au moment où les feuilles sont encore jeunes, c'est-à-dire avant la fermeture de la pomme. Les plantes hôtes sont les premiers indicateurs de l'infestation d'une parcelle. En effet, c'est en observant les dégâts foliaires qu'une première estimation peut être faite. C'est pour cela qu'il est important de surveiller particulièrement les premières semaines du

3). Les parasitoïdes de *Plutella xylostella*



Deux types de parasitoïdes peuvent entraîner la mort des individus de *Plutella xylostella*. Ils se distinguent par leur mode d'infestation.

Le premier groupe se compose des espèces *Oomyzus skolowskii* (figure 6.A) et *Brachymeria sp.* (figure 6.B). Les adultes

Figure 6 : Principaux parasitoïdes de *Plutella xylostella*, A: *Oomyzus skolowskii*, B: *Brachymeria sp.*, C: *Apanteles litae* et D: *Cotesia plutellae*

de ces deux espèces pondent leurs œufs alors que l'individu de *Plutella xylostella* est encore au stade larvaire. Ainsi, les parasitoïdes forment leur cocons à l'intérieur de la nymphe de *Plutella xylostella*, la détruisent complètement et émergent du cocon de cette dernière. On parle ici d'un cycle de développement endogène.

Le second groupe de parasitoïdes se compose des espèces *Apanteles litae* (figure 6.C) et *Cotesia plutellae* (figure 6.D). Les adultes de ces deux espèces pondent leurs œufs à l'intérieur des cocons de *Plutella xylostella*. Ainsi, le développement des parasitoïdes entraînent la mort de la nymphe mais la trace de sa présence persiste dans le cocon. Après le développement larvaire, les parasites forment leur cocon contre le cocon de *Plutella xylostella*. Ce nouveau cocon peut se situer à l'intérieur (*Apanteles litae*) à l'extérieur (*Cotesia plutellae*) du cocon de l'hôte (Ayalew, 2011).

Ainsi, de nombreuses espèces permettent de réguler les populations de ravageurs responsables des dégâts des parcelles de choux.

Une fois le volet biologique du sujet approfondi, il est nécessaire de présenter les connaissances existantes sur le lien entre le paysage et la régulation des insectes ravageurs des cultures.

2. Connaissances sur le paysage

Le contrôle des ravageurs des cultures tend de plus en plus à considérer le paysage comme un élément important et comme niveau d'organisation spatiale pour appréhender la dynamique des populations d'insectes. Ce concept est bien défini par le terme d'areawide-IPM (Integrated Pest management) qui regroupe à la fois les connaissances sur les liens entre la biologie des ravageurs, le paysage et les pratiques agricoles.

Peu d'études intègrent à la fois les pratiques agricoles et l'effet du paysage pour expliquer la présence et l'abondance des insectes ravageurs, même si de nombreuses études expliquent l'abondance des ravageurs à la lumière d'un de ces deux paramètres.

De nombreuses études montrent l'effet de la composition et la forme du paysage sur les paramètres d'infestations des cultures à l'échelle du paysage ou à des échelles plus petites (Chaplin-Kramer and Kremen, 2012). Il a été montré que la réponse des ravageurs à la complexité du paysage dépend de l'espèce d'insectes (Chaplin-Kramer et al., 2011) ou des espèces n'appartenant pas à la classe des insectes (Thies et al., 2003) qui est considérée. Ainsi le lien entre le taux d'infestation des différents ravageurs et la composition du paysage a été largement étudié (Veres et al., 2013), en parallèle de la recherche du lien entre l'abondance des ravageurs et la disposition des plantes hôtes dans le paysage. (Wilkinson and Landis, 2005). La caractérisation de l'influence de la complexité du paysage sur les parasitoïdes a également été mise en évidence (Gols et al., 2005), de même que l'enjeu de la gestion des espèces aux vues des services écosystémiques rendus (Kremen and Chaplin-Kramer, 2007).

Un ensemble d'études se sont concentrées sur la caractérisation du paysage. Il apparaît alors que l'indice de Shannon est le plus apte à représenter cette caractérisation (Spellerberg and Fedor, 2003).

Enfin, la gestion du paysage ne peut se faire sans impacter plus ou moins fortement les pratiques culturales qui y sont mis en place.

3. Connaissances sur les pratiques agricoles

Les pratiques agricoles sont également observées afin d'améliorer la lutte intégrée contre ces ravageurs. Dans le cas de la lutte chimique, la problématique est complexe puisque les insectes acquièrent au fil des générations des caractères de résistance. Ainsi, quelques études ont pu mettre en évidence des réponses particulières de l'espèce *P.*

xylostella à des insecticides comme le régulateur de croissance novaluron (Ayalew, 2011) ou l'extrait de neem (Charleston et al., 2006).

Certaines études s'intéressent au lien existant entre l'intensification de l'agriculture et la gestion des ravageurs (Jeger, 1999) et tentent de faire le lien avec les parasitoïdes et les effets sur le paysage. Ainsi, une étude de 1985 établit le lien existant entre le contrôle de *Plutella xylostella* et d'autres insectes ravageurs, le système d'irrigation par rampes, un insecticide et certains agents de contrôle biologiques (Nakahara et al., 1986).

Enfin, un article a été publié en 2005 afin de faire la synthèse des connaissances existantes sur le sujet de la régulation de la teigne du chou (Sarfraz et al., 2005). Cette article rassemble les conclusions de plus de 150 articles différents et regroupe à la fois les aspects biologiques concernant *Plutella xylostella* mais également ses principaux parasitoïdes et prédateurs et sur les pratiques de gestions intégrés de ces populations.

4. Synthèse des connaissances

Certaines études tentent de mettre en évidence le lien existant entre les pratiques culturales, le paysage et la gestion des ravageurs. Ces articles sont peu nombreux et s'appliquent dans certains cas à des problématiques générales (Rusch et al., 2010), à une culture donnée comme le colza (Furlong et al., 2008), l'amandier (Eilers and Klein, 2009) ou encore le cerisier (Stutz and Entling, 2011) ou à une aire de répartition géographique ciblée comme l'Ethiopie (Ayalew and OgoI, 2006) ou les Pays-Bas (Bianchi et al., 2008).

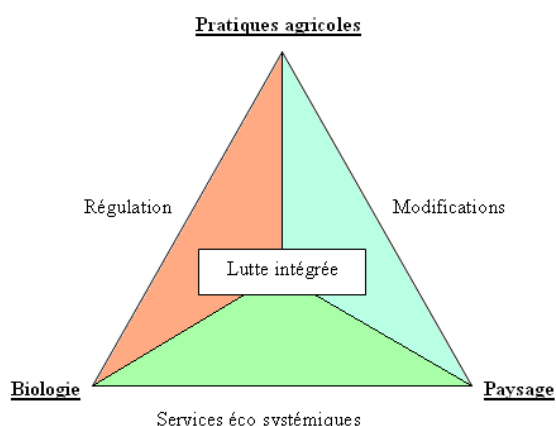


Figure 7 : Schéma récapitulatif des connaissances

hypotheses établissent que certaines variables environnementales et/ou culturelles ont un effet sur la dynamique des populations du ravageur ciblé et sur sa régulation naturelle.

B. Problématiques

Ce stage s'inscrit dans une problématique large concernant l'intensification écologique à travers la conception de systèmes de cultures moins dépendants des pesticides. Selon la définition de l'axe prioritaire de recherche du CIRAD⁵, 'l'intensification écologique, c'est concevoir des systèmes de production plus productifs, durables, économes en intrants et moins nocifs pour l'environnement. C'est aussi créer des variétés mieux adaptées à leur milieu, inventer de nouvelles techniques de lutte contre les maladies et les ravageurs. C'est comprendre le fonctionnement de la nature pour exploiter ses ressources sans la détruire, produire plus et rompre avec les pratiques fondées sur l'utilisation excessive et massive de pesticides, d'engrais chimiques, d'eau et d'énergie fossile'.

⁵ <http://www.cirad.fr/nos-recherches/axes-prioritaires/intensification-ecologique/questions-de-recherche>

L'intensification écologique lie ainsi plusieurs domaines de compétences apportant chacun certains enjeux, qu'ils soient d'ordre agronomiques (système plus productifs), sanitaires (moins nocifs), environnementaux (durables) et économiques (économes en intrants).

Ce stage tente donc de mettre en lumière les facteurs agronomiques (pratiques culturales) et paysagers (environnementaux) qui influencent la régulation des insectes ravageurs des cultures de chou dans la zone des Niayes au Sénégal.

C. Déroulement du stage

Comme l'indique le sujet du stage, l'étude porte sur la recherche des liens existants entre les pratiques agricoles, le paysage et la régulation naturelle des insectes ravageurs des cultures dans la zone des Niayes. Il faut donc tout d'abord collecter ces différents indices afin d'en analyser les liens.

La collecte des différentes données nécessaires à la mise en place des variables de l'analyse peut être découpée en trois temps selon le type de variables considérée. Il faut tout d'abord veiller à collecter les données relatives aux indicateurs biologiques. Pour cela, il est nécessaire de mettre en place un protocole cohérent pour l'échantillonnage des parcelles, la collecte des individus sur ces parcelles et le suivi de ces individus en laboratoire. Ensuite, il est nécessaire de collecter les données relatives aux pratiques culturales auprès des agriculteurs en vérifiant au maximum les informations fournies. La dernière phase de la

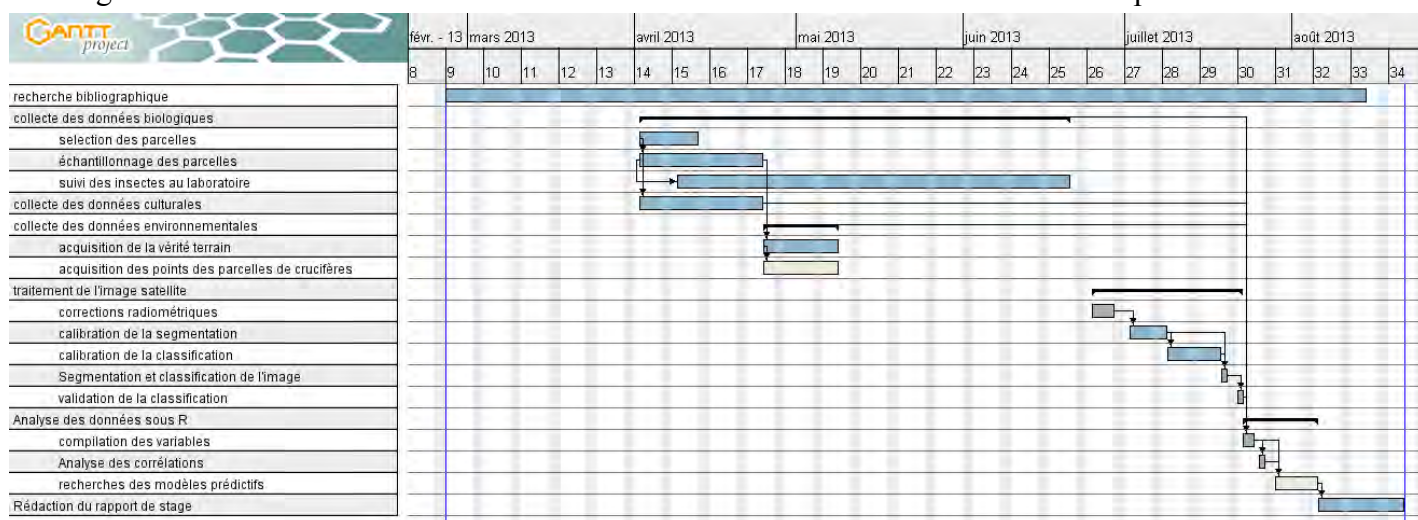


Figure 8 : Diagramme de Gantt du stage

collecte est celle des données environnementales pour l'étude de l'effet du paysage sur la régulation naturelle des insectes ravageurs des cultures de choux. Ici, il n'est pas possible de collecter exhaustivement ces informations environnementales, c'est pourquoi nous avons fait le choix d'utiliser une image satellite THRS (Très haute résolution spatiale) afin de réaliser une carte d'occupation du sol de toute la zone d'étude. Pour la réalisation de cette carte, une classification assistée par ordinateur a été conçue à partir d'une image satellite. Il a alors été nécessaire de s'assurer de sa viabilité afin en assurer la fiabilité. Une fois les données collectées, il faut les analyser de manière statistique afin de déterminer les modèles qui expliquent les variables biologiques.

Comme le montre le diagramme de Gantt (figure 8), la résolution de ces différents points de problématique sert de trame chronologique au stage et au rapport.

On retrouve aisément les trois blocs d'actions du projet que sont les phases de collectes des données (biologiques, culturales et environnementales), le traitement de l'image et l'analyse des données. Les phases de collectes sont clairement séparées des deux autres. Ceci est dû à la localisation du stage à ce moment précis. En effet, la série de collecte s'est

effectuée au Sénégal, dans la région de Niayes, et le traitement de l'image ainsi que l'analyse des données se sont déroulés à mon retour de mission à Montpellier.

Une fois les connaissances existantes spécifiées et les enjeux de l'étude bien déterminés, il est important de mieux connaître la zone sur laquelle se déroulera l'étude.

D. Description de la zone d'étude

Même si la connaissance du socle institutionnel et intellectuel est importante pour comprendre l'étude qui sera menée, il est tout aussi primordial de connaître et comprendre les caractéristiques de la zone d'étude.

1. Les Niayes en général

La connaissance de la zone d'étude est essentielle afin de mieux cerner les contraintes imposées au système et ses caractéristiques. Il est nécessaire de la localiser afin de pouvoir faire le lien avec les caractéristiques physiques et le climat. Il est ensuite intéressant de se renseigner sur les différentes utilisations de sols présents sur la zone.

Situé dans la partie ouest de l'Afrique, le Sénégal présente un climat semi-désertique (moitié nord plus aride) qui se traduit par l'alternance d'une saison sèche (de fin octobre à juin) et d'une saison humide. La période pendant laquelle s'est déroulée la phase d'expérimentation était donc une saison sèche où les apports en eau par la pluie sont nuls.

La zone des Niayes est une bande de terre en bordure de littoral sénégalais s'étendant entre Dakar et Saint Louis. (Voir figure 9). Cette bande étroite s'étend sur une longueur de 180 km et une largeur variant entre 25 et 30 km.

Dans cette zone, le sol est majoritairement sableux avec une nappe phréatique peu profonde, ce qui explique en partie le développement des activités maraîchères en saison sèche malgré le manque d'eau pluviale. Cette bande s'étend d'une part entre la presqu'île du Cap-Vert (occupée par la ville de Dakar au sud) et la frontière mauritanienne (au nord). De plus, 80 % de la production horticole sénégalaise (fruits et légumes) sont concentrés dans les Niayes même si l'aviculture industrielle y est très présente.



Figure 9 : Localisation de la zone des Niayes

2. La zone d'étude en particulier

La zone d'étude est située dans la partie Sud des Niayes et s'étend sur une surface de 375km² environ, à une cinquantaine de kilomètres de Dakar.

En plus du Lac Rose indiqué sur la carte (Lac Retba), deux lacs asséchés en saison sèche sont présents sur la zone. Il s'agit du lac Touma et du Lac Mbawane (figure 10). La présence de ces lacs temporairement secs modifie le paysage en y



Figure 10 : Localisation de la zone d'étude (en vert, source : Word Street Map)

ajoutant des espaces dépressionnaires et sédimentaires plutôt propices aux cultures.

Une fois les problématiques et les enjeux du stage clairement énoncés, il est temps de s'intéresser aux matériels utilisés pour la réalisation des différentes étapes du stage.

III. Matériels

Cette partie décrit l'ensemble des données et des outils qui ont été utilisés au cours de ce stage. Dans un premier temps, il est nécessaire de présenter les caractéristiques de l'image satellite et les logiciels utilisés.

A. Caractéristiques des images satellites Pléiade

La connaissance des caractéristiques des images utilisées est nécessaire pour comprendre et de régler les problèmes relatifs à l'hétérogénéité de l'information. Les jeux d'images utilisés pour cette étude sont issus du réseau Pléiades.

1. Présentation des images satellites Pléiades

Le réseau Pléiades est une constellation de deux satellites déployés par Astrium⁶. Les satellites Pléiades 1A et Pléiades 1B opèrent à la même orbite avec un positionnement à 180° l'un de l'autre. Les deux satellites jumeaux délivrent des images THRS (Très Haute Résolution Spatiale) avec une capacité d'acquisition d'un jour sur l'ensemble des points de la planète.

Les caractéristiques des images Pléiades sont identiques pour toutes les images produites. Pour les images panchromatiques, la résolution spatiale est 50 cm avec une couverture spectrale de 470 à 830 nm. Pour les images multi-spectrales, la résolution est de 2m et est composée de quatre bandes spectrales : le bleu (430 à 550 nm), le vert (500 à 620 nm), le rouge (590 à 710 nm) et le proche infrarouge (740 à 940 nm)

En termes de formats d'images, les produits Pléiades peuvent être délivrés aux formats JPEG2000 (avec compression normale ou optimisée) ou au format geoTIFF (avec un encodage de 16 bits). C'est dans ce dernier format que sont présentées les données de l'étude. Le système de coordonnées utilisé pour les images satellites et les données dérivées est le WGS84 28N UTM. WGS84 (World Geodetic System 1984) est un système géodésique très répandu et utilisé notamment pour les mesures GPS. C'est une manière de représenter l'ellipsoïde terrestre. La Transverse universelle de Mercator UTM (Universal Transverse Mercator) est un type de projection permettant de localiser les points sur cet ellipsoïde. Enfin, le terme 28N permet de savoir dans quelle partie de l'hémisphère sud on se trouve.

La connaissance des caractéristiques des images est nécessaire afin de rendre standard les protocoles utilisés par la suite

2. Caractéristiques des images Pléiades utilisées

Deux lots d'images ont été livrés afin de mener cette étude. Le fait, que l'on ait reçu l'image en plusieurs morceaux est essentiellement lié aux poids des images qu'il est possible de livrer en une fois. Les caractéristiques physiques de ces quatre images sont regroupées dans le tableau 1. Chaque lot d'images a reçu un numéro d'identification (repris dans le nom du dossier).

⁶ Site internet d'Astrium

Tableau 1 : Caractéristiques physiques des images Pléiades

Image	Multi 3303	Panchro 3303	Multi 3304	Panchro 3304
heure d'acquisition	16 :02 :24.782	16 :02 :24.844	16 :04 :49.964	16 :04 :56.581
Angle d'acquisition	24.82°	24.82°	19.25°	19.25°
X top left	254 664	254 666	262 166	262 141
Y top left	1 651 844	1 651 842	1 659 062	1 659 060
X bottom right	271 086	271 085	282 322	282 321
Y bottom right	1 632 118	1 632 119	1 633 554	1 633 555

La date d'acquisition de l'image a été fixée au 16 février puisque cette période est propice à la détection des périmètres maraichers. La connaissance de la date, de l'heure exacte et de l'angle d'acquisition des images permettra par la suite d'effectuer les prétraitements (corrections radiométriques et géométriques) nécessaires.

Des données relatives au dispositif de prise de vue (l'objectif du satellite Pléiades 1A) sont également nécessaires pour effectuer les corrections radiométriques nécessaires.

Tableau 2 : Caractéristiques spectrométriques des images Pléiades

Bandes	Rouge	vert	Bleu	Proche infra rouge	Panchro
Gain	10.33	9.16	9.59	15.64	11.36
Biais	0	0	0	0	0
Irradiance solaire	15944	1830	1915	1060	1548

La phase de traitement des images satellites a été l'occasion d'utiliser certains logiciels spécifiques qu'il est intéressant de détailler ici.

B. Logiciels

Les logiciels utilisés pour mener à bien cette étude ont servis à réaliser deux types de tâches qui permettent de les distinguer.

1. Logiciels de traitements de l'image

L'ensemble des prétraitements et des de l'image satellite Pléiades ont nécessité la prise en main de plusieurs logiciels.

Le logiciel ERDAS⁷ Imagine (version 2010) a été utilisé pour l'ensemble des prétraitements radiométriques et géométriques de l'image grâce à la fonction model maker qui permet de coder des algorithmes de prétraitements adéquats. Il a également permis la réalisation de mosaïque à 5 bandes à partir des deux images multi spectrales et des deux images panchromatiques. De plus, le logiciel ERDAS a été utilisé pour extraire une zone plus petite sur laquelle ont été développées les chaînes de traitement (segmentation et classification d'image) sous le logiciel e-cognition.

C'est à partir de ce groupement d'images caractéristiques des paysages qu'a été testée la segmentation puis la classification dans le logiciel e-cognition⁸ dans sa version 8.7. Ce logiciel a permis de réaliser l'arbre de classification et de valider

**Figure 11 : Logo ERDAS****Figure 12 : Logo E cognition**

⁷ <http://geospatial.intergraph.com/Homepage.aspx>

⁸ <http://www.ecognition.com/>

l'ensemble des paramètres et des seuils de séparations des classes d'occupation du sol. Une fois les algorithmes de segmentation et de classification développés, l'image entière a été traitée via un serveur externe permettant de lancer des calculs lourds. A l'issue des traitements, nous avons obtenu une carte d'occupation du sol en 13 classes au format shape, reconnu par les logiciels ESRI.

Enfin, pour le calcul de variables environnementales dérivées de la carte d'occupation du sol, nous avons utilisé le logiciel SIG ArcGIS⁹ (version 10.0), et plus spécifiquement le module 'Patch Analyst' dédiés aux calculs d'indices paysagers complexes. Ce logiciel a également permis de créer l'intégralité des couches thématiques liées à la classification (mer, verger, végétation naturelle...)



Figure 13 : Logo ArcGIS

En parallèle des logiciels de traitements de l'image satellite, un ensemble d'outils informatiques ont été utilisés afin d'acquérir, de gérer et d'analyser les données alphanumériques recueillis grâce à l'image et sur le terrain.

Figure 16 : Logo R

2. Logiciels d'acquisition, de gestion et d'analyse des données

Le premier logiciel utilisé pour l'acquisition, la gestion et l'analyse des données est le logiciel ArcPad¹⁰ dans sa version 8.0. Appartenant à la gamme ESRI, il est totalement compatible et intégrable au logiciel ArcGIS. Après la création d'un projet spécifique sous ArcPad à partir des couches géo référencées du logiciel ArcGIS, il est possible d'utiliser les informations du projet de manière nomade et de les mettre à jour directement sur le terrain. Ce logiciel a ainsi permis la collecte de l'intégralité des données relatives aux pratiques culturelles en même temps que la collecte des données biologiques.



Figure 14 : Logo ArcPad

Le deuxième logiciel mis à contribution dans ce cadre est le logiciel Access. Il appartient à la suite bureautique de Microsoft dans sa version 2007¹¹. Ce logiciel permet de gérer l'ensemble des données alphanumériques collectées sur le terrain et au laboratoire. La création de différentes tables reliées entre elles permet une saisie et une lecture plus facile des données. Ceci permet aussi la création de requêtes plus ou moins complexes. Ces différentes tables et requêtes ont permis de calculer les variables biologiques utilisées dans l'étude statistique.



Figure 15 : Logo Access

La dernière étape dans la gestion, acquisition et analyse des données est la phase d'analyse des données. Le logiciel utilisé pour ce traitement est le logiciel R¹² (version 2.1). Il s'agit d'un logiciel en libre accès. Grâce au chargement de nombreuses librairies, ce logiciel permet l'utilisation de fonctions statistiques et graphiques à la fois complexes et performantes. Il présente donc une grande robustesse de calcul malgré une interface peu conviviale au premier abord.



Figure 16 : Logo R

La connaissance des outils utilisés pour la réalisation du stage permet de comprendre les contraintes techniques pesant sur la méthodologie.

IV. Démarche méthodologique

Trois étapes ont été nécessaires pour répondre à la question de l'effet des pratiques agricoles et du paysage sur la répartition spatiale et l'abondance des insectes ravageurs ciblés et de leurs ennemis naturels : une première étape de collecte d'informations de terrain et d'analyse en laboratoire, une seconde étape dédiée à la réalisation d'une carte d'occupation du

⁹ <http://www.arcgis.com/features/>

¹⁰ <http://www.esrifrance.fr/arcpad.aspx>

¹¹ <http://office.microsoft.com/fr-fr/access/>

¹² <http://www.r-project.org/>

sol dérivée d'une image satellite Pléiades et enfin l'analyse statistiques des relations entre les données sur les insectes avec les données sur les pratiques agricoles et du paysage.

A. Phase de terrain et laboratoire

Lors de cette phase, trois grandes étapes ont été recensées et chacune a nécessité la mise en place d'un protocole. Ces étapes sont : la sélection des parcelles, le comptage et la collecte des insectes au champ et enfin le suivi des individus au laboratoire.

1. Choix des parcelles

Le choix des parcelles est essentiel puisque c'est sur cette unité qu'est basée l'intégralité de l'étude. La sélection des parcelles s'est déroulée durant le mois d'avril 2013. La zone d'étude a été divisée en trois blocs, permettant ainsi un quadrillage plus efficace de la surface à traiter dans la journée. Lors de la première phase de la collecte (du 8 au 11 avril), les parcelles qui ont été choisies ont été soit observées, soit juste géo référencées car encore trop récemment repiquées (moins de 15 jours après le repiquage) afin d'être connues et observées lors de la deuxième phase. Le choix des parcelles a été conditionné par la distance existant entre la parcelle candidate et l'ensemble des parcelles déjà sélectionnées. En effet, l'étude environnementale de ces parcelles étant réalisée dans un rayon de 1000 m autour de ces parcelles, une distance minimale de 2 km entre deux parcelles a été observé afin d'éviter les phénomènes d'auto-corrélation entre les variables.

La géo localisation des parcelles s'est faite à l'aide d'un GPS de la gamme Trimble Juno 3B. Cet outil est capable de supporter le logiciel ArcPad et permet donc de visualiser les données sur le terrain et de les mettre à jour. De plus, l'écran large ainsi que le processeur puissant permet la visualisation de cartes de grande taille et de précision importante.

Malgré la puissance et la précision du GPS Trimble, la navigation en temps réel avec le calcul des distance entre les différentes parcelles est compliquée à mettre en place et souvent trop longue. C'est pour cela que, lors de chaque sortie, un GPS Garmin de la gamme Oregon a été joint au matériel emporté. Ce GPS a été créé à l'origine pour les randonnées pédestres et les déplacements motorisés. Il permet d'entrer une série de points de repères (dans le cas de l'étude, il s'agit du point central des parcelles) et de calculer la distance à vol d'oiseau entre la position du GPS et l'ensemble des points enregistrés. Ceci permet une navigation aisée et le respect d'une distance minimale entre deux parcelles. Le choix d'une parcelle plutôt qu'une autre a été motivé par l'accessibilité aux informations sur les pratiques culturales auprès du propriétaire de la parcelle. Ainsi, il a été possible de collecter les données relatives aux paramètres suivants : la densité (ou le nombre de plant par m² de sol), le nombre de traitements insecticides effectués avant et après le repiquage et le produit utilisé, le type d'irrigation mis en place et la variété plantée.

Les autres variables culturales ont été obtenues après regroupement et observation de la parcelle. C'est le cas pour la note de recouvrement du sol par les adventices. La figure 17 montre le barème permettant d'attribuer une note adventice en fonction du pourcentage de surface recouvert par d'autres plantes que le chou.



Figure 17 : Barème d'attribution de la note d'adventice en fonction du pourcentage de surface recouverte par des adventices

Une note de sélectivité a été attribuée aux traitements insecticides (voir annexe 1: exemple de calcul de la sélectivité). Ce calcul en fait donc un indice composite combinant le spectre d'action des produits utilisés et le nombre de traitements effectués. Une note élevée

signifie une faible sélectivité des traitements insecticides effectués, et donc un impact négatif probable sur les ennemis naturels.

Une fois les parcelles sélectionnées, l'étape d'échantillonnage et de capture des insectes peut être effectuée.

2. Capture des insectes

Le comptage et la collecte des insectes sur les parcelles suivies est la première étape de récolte des données biologiques qui permettront l'analyse ultérieure. Quatre phases de collecte espacées de 2 semaines ont été menées du 6 avril au 5 juin 2013. Chaque phase de collecte s'est faite sur trois jours à raison d'environ 10 parcelles par jour.

Pour chaque parcelle, trois observations ont été faites au total à deux semaines d'intervalle. A chaque observation, 30 plants ont été sélectionnés et observés. La méthode d'échantillonnage retenue est le lancer de règle (voir figure 18) en choisissant le plant le plus proche du point de chute de la règle, après avoir divisé la parcelle en bande de taille variable selon le nombre d'observateurs (entre 2 et 4).

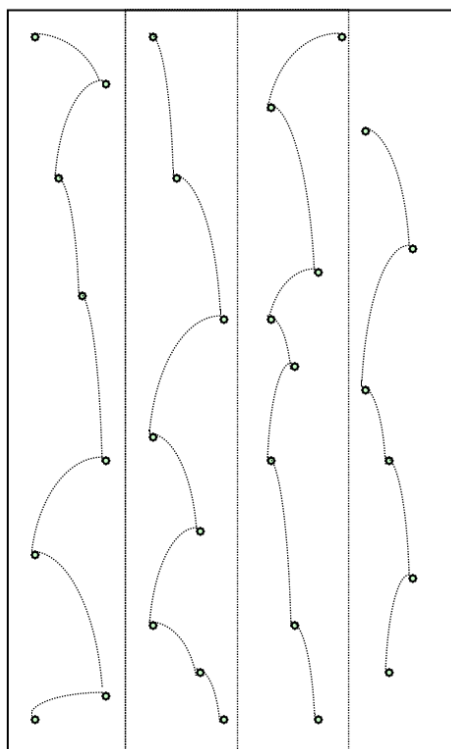


Figure 18 : Exemple de protocole de choix des plants à échantillonner (sur 4 bandes)

3. Suivi en laboratoire

Au laboratoire, les nymphes issues des larves collectées sont placées dans des tubes Eppendorf et des piluliers (phases 1 et 2) ou des boîtes Nunclon 24 (phases 3 et 4). La date de nymphose est notée dans la grille de suivi et chaque individu reçoit un code d'identification sous la forme [a.b] avec a : le numéro de la parcelle et b : Le numéro de l'individu dans la grille de suivi. Il est alors facile d'avoir une vue d'ensemble et un contrôle clair sur l'ensemble des individus et leur données.

Les individus morts au stade larvaire reçoivent également un code d'identification avec la saisie de la date de mortalité. Un extrait de fiche de suivi des individus au laboratoire est présenté dans le tableau 3.

Tableau 3 : Fiche de suivi des individus de *Plutella xylostella* en laboratoire

Num parcelle	num obs.	Num individu	stade à la collecte	date mortalité larve	date nymphe	date adulte	sexe de l'adulte	Date parasite	nombre de parasite	espèce	sexe parasite	Stade	contenant
1	1	32	larve	14/04/13								larve	
3	3	7	larve	08/05/13								larve	
4	1	18	larve		10/04/13			16/04/13	1	Apanteles	Femelle	parasite	
6	3	13	nymphe			09/05/13	Male					adulte	pillulier

Le suivi des nymphes s'effectue durant les 21 jours suivant la collecte. A la fin de cette période, les nymphes exemptes d'émergence sont notifiées avant d'être éliminées. Deux types d'émergence peuvent survenir à partir du stade nymphe, tout d'abord l'apparition d'un adulte de *P. xylostella* (il est alors noté la date d'émergence et le sexe de l'individu), et ensuite l'apparition d'un parasitoïde (il est fait mention de la date d'émergence, de l'espèce et du sexe du parasitoïde). Dans le cas d'Oomyzus, il est possible de rencontrer plusieurs individus pour la même émergence.

Les adultes et les parasitoïdes seront ensuite conservés dans de l'alcool à 90° avec mention de la date de collecte et du numéro de parcelle afin de permettre une étude ultérieure de génétique des populations.

Dans le fichier regroupant l'ensemble des informations de suivi est ajouté, par individu, le stade à la fin de l'observation (Larve, nymphe, adulte ou parasitoïde) ainsi que le contenant ayant servi à la séparation des nymphes (boîte Nunclon, pillulier, tube Eppendorf).

B. Traitement de l'image

Le traitement des images satellites a été réalisé en plusieurs étapes. Une première phase de prétraitement de l'image satellite, une phase de segmentation et de classification d'image et enfin une phase d'évaluation de la qualité des résultats de la carte d'occupation du sol obtenue.

1. Prétraitements des images satellites

La première étape des prétraitements des images satellites est l'application d'un ensemble de corrections radiométriques et géométriques afin de convertir le signal enregistré par le satellite en une variable physique mesurable. Pour cela, on procède à des corrections géométriques et radiométriques sur chaque bande pour pouvoir comparer les valeurs des pixels les uns aux autres et dans le soucis d'une répliquabilité des algorithmes de segmentation et de classification.

1). Correction radiométriques

Les images satellites utilisées pour la classification sont issues des produits Pléiades et représente les compte numérique de chaque pixel pour chacune des bandes acquises. La conversion de ces comptes numériques en luminance puis en réflectance permet de standardiser les valeurs de chaque pixel afin de rendre la méthode la plus reproductible possible.

La conversion des comptes numériques en luminance permet de reconstituer les valeurs de luminance à l'entrée du capteur satellite et ainsi de réduire les effets de l'atmosphère. Pour cela, l'algorithme présenté en annexe 2 a été mis en place avec les paramètres présentés dans le tableau 2. Cette transformation consiste à créer un indice correspondant à la valeur du compte numérique divisé par le gain absolu de la bande (identique pour chaque pixel de l'image) et y ajouter le biais de calibration. Cette manipulation permet donc d'effacer les variations existantes entre deux appareils de prise de vue (ou deux instants sur le même appareil) en termes de calibration de caméra.

De plus, l'accès à la valeur de luminance pour chaque pixel a permis ensuite de calculer la valeur de réflectance pour chaque pixel. Pour cela, l'algorithme présenté en annexe 3 a été mis en place avec les paramètres présentés dans les tableaux 1 et 2. Cette transformation permet de s'affranchir des variations dues à la puissance du rayonnement solaire (différente entre les mois d'hiver et d'été et entre le matin et le midi) et à la position du soleil dans le ciel (lumière rasante ou zénithale). Ainsi, des images prises à des moments différents pourront être comparées et traitées de manière similaire.

Une fois les étapes de standardisation de chaque bande de l'image effectuée, une mosaïque des deux blocs de l'image a été réalisée afin de faciliter la suite des traitements

2). Préparation des données

La préparation des données a permis de créer une image utilisable pour les tests de classification et de segmentation ainsi qu'une image utilisée pour le traitement total de la zone d'étude.

Tout d'abord, les deux images ont été redécoupées en 25 dalles chacune. Sur ces 50 dalles, 4 ont été sélectionnées comme représentatives des différents paysages observables et regroupées en une bande composite sans géo référencement. Cette bande a servie par la suite

à la mise au point de la segmentation et de la classification à partir d'une imagerie représentative de l'image totale mais plus légère et donc plus maniable. Les zones utilisées pour la création de cette image test sont représentées sur la figure 19.A par les rectangles rouges. L'image obtenue à la suite de ce traitement est également présentée dans la figure 19.B

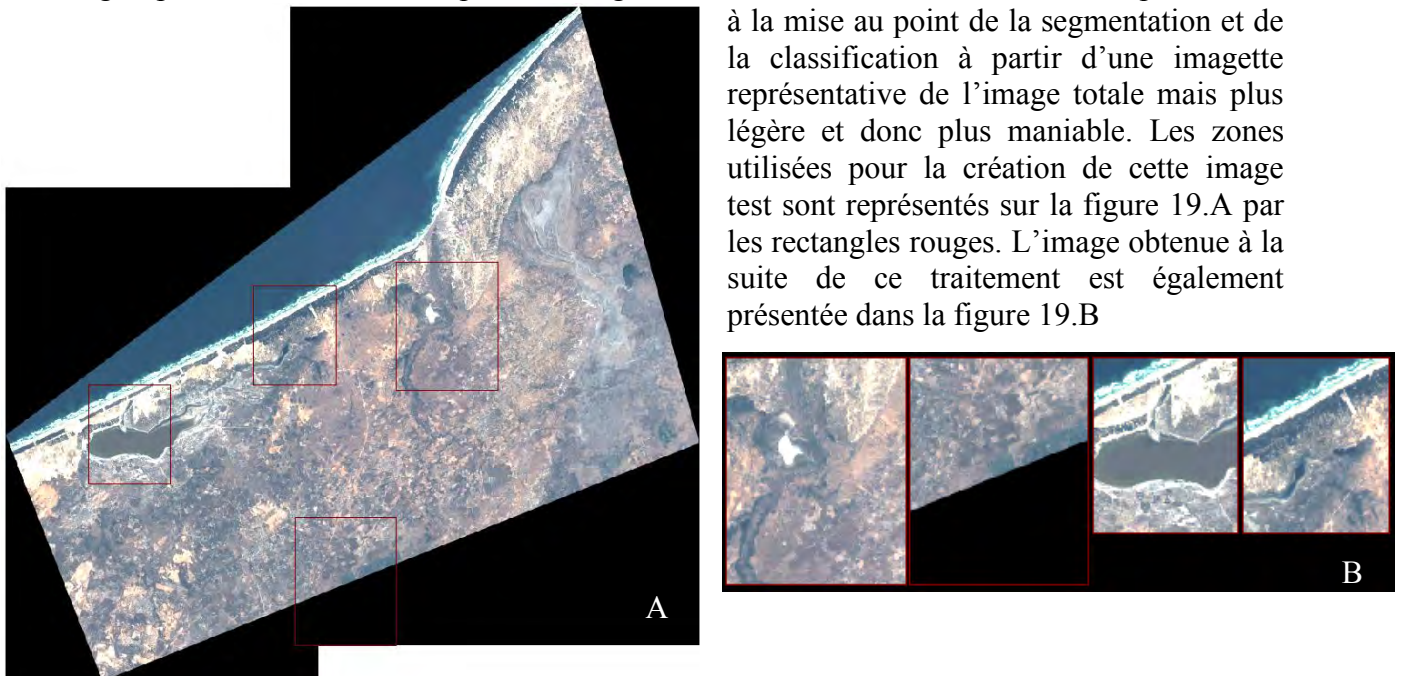


Figure 19 : Zones utilisées pour la création de l'image test (A) et image test obtenue (B)

Enfin, les deux images multi spectrales ont été fusionnées avec leur image panchromatique respective avant d'être regroupées dans une unique mosaïque. Il a ainsi été créé une image possédant 5 bandes et couvrant l'intégralité de la zone d'étude. Ceci a permis un traitement unique de l'intégralité de la couche grâce au serveur de calculs.

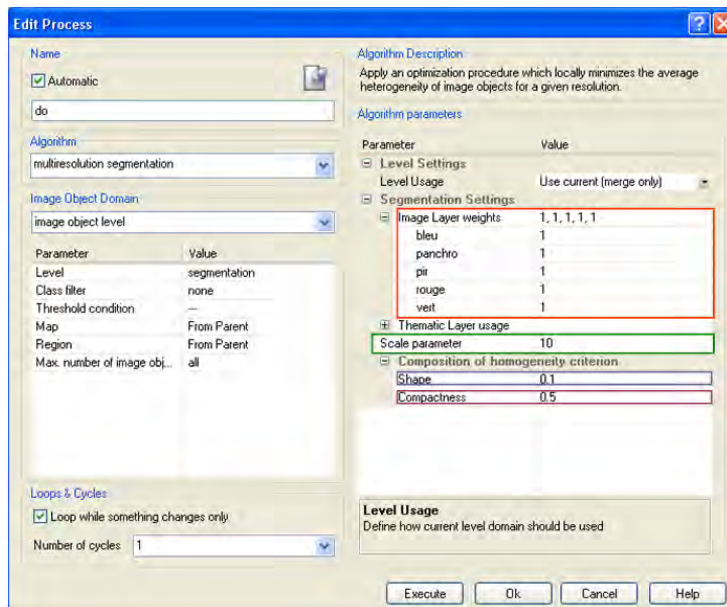
Une fois les algorithmes de segmentation et de classification développés, la chaîne de traitement a été réalisée sur la totalité de l'image satellite et ses 5 bandes spectrales (B, V, R, PIR et Panchromatique). Les temps de calcul étant très long, cette opération a été réalisée via un serveur de calcul externe.

2. Segmentation et classification de l'image

L'ensemble des étapes de segmentation et de classification d'image ont été développées sur l'image test (Cf partie IV. B. 2) Il sera ici présenté la démarche et les paramètres retenus pour la segmentation de l'image puis pour la classification.

1). Segmentation

En traitement des images, la segmentation est la division de cette image en objets homogènes selon un ou plusieurs critères déterminés : couleur, texture, indice radiométrique.... Le niveau de segmentation permet ainsi de discriminer des objets proches (en termes de radiométrie ou de formes) mais de nature différente. La séparation d'objets plus ou moins semblables dépendra de la finesse de la segmentation. Ici, on utilise le logiciel e-cognition.



Une fois l'image chargée dans le logiciel, il est simple de créer une procédure (figure 20) permettant la segmentation de l'image. Pour cela, il faut entrer quatre paramètres afin de caractériser la segmentation. Chacun de ces paramètres est représentés par un bloc de couleur différente dans la figure 20.

Le premier de ces paramètres est le poids de chacune des bandes de l'image (en rouge). Par défaut, chaque bande reçoit une valeur de 1. Plus le poids est élevé, plus la valeur de cette bande aura un impact important dans le choix de la meilleure segmentation.

Le deuxième paramètre pour la segmentation est le 'scale parameter', le paramètre d'échelle (en vert). Il définit

Figure 20 : Boîte de dialogue du logiciel e-cognition pour le choix des paramètres de segmentation

une limite sans dimension qui détermine l'hétérogénéité maximale permise par les objets. Ainsi, les objets résultant d'un même paramètre d'échelle seront plus petits pour des données hétérogènes et plus grands pour des données homogènes.

Les troisième et quatrième paramètres entrent dans le groupe des paramètres d'homogénéité de composition. Il s'agit du paramètre de forme ('shape' en bleu) et du paramètre de compacité ('compactness' en violet). Le 'shape' (i.e la forme contre la couleur) correspond à l'homogénéité spectrale de l'objet. Une pondération du critère 'shape' à 0 implique que les objets seront optimisés uniquement par leur homogénéité spectrale. Pour le paramètre de 'compactness', la forme est définie par deux critères sous-jacents : le 'smoothness' et le 'compactness'. Une pondération de critère de compacité à 0 implique que les objets auront un contour très découpé. L'utilisateur peut donc agir à la fois sur les critères d'homogénéité et d'hétérogénéité (figure 20) en fixant des seuils aux différents paramètres.

Il faut donc trouver la combinaison de valeurs la plus adaptée au niveau de discrimination recherchée afin de séparer de manière adéquate les différents objets. La recherche des paramètres de segmentations optimaux est une tâche plutôt longue puisqu'entièrement manuelle. Les résultats des différents tests effectués pour trouver les bons paramètres sont présentés en annexe 4 avec les observations pour chaque modalité.

Les paramètres de segmentation suivants ont été retenus et ont permis de réaliser la segmentation présentée en figure 21:

- Poids des bandes :
- Rouge : 1
 - Vert : 1
 - Bleu 1
 - Proche infra rouge : 5

- Panchromatique : 5
 Paramètre d'échelle : 50
 Forme : 0.1
 Compacité : 0.9

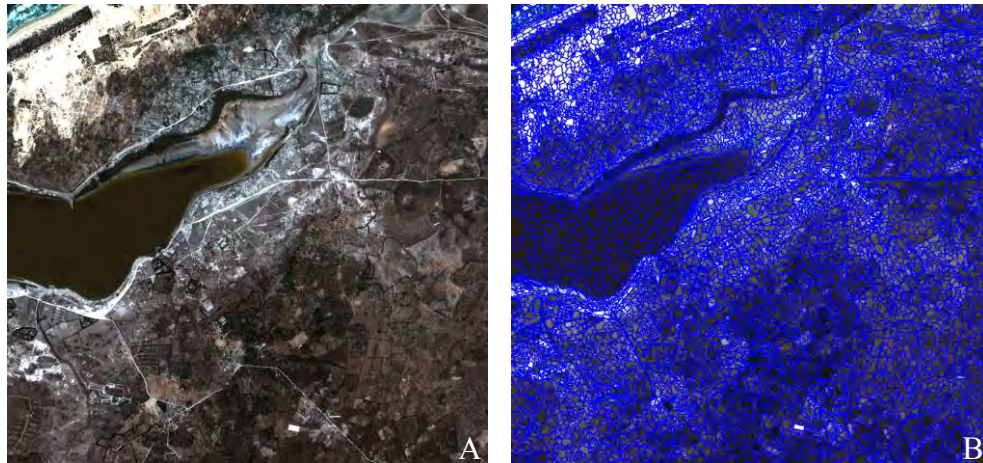


Figure 21 : Résultats de la segmentation (A : image brute et B : image segmentée)
 Après la phase de segmentation, il a été nécessaire de réaliser de nombreux tests et ajustement afin de classifier l'image.

2). Classification

Tout comme la segmentation, la phase de mise au point des algorithmes de classification a été réalisée sur l'image test. Le choix du nombre de classes et de leur caractéristiques ont nécessité une phase d'ajustement longue à travers la mise en place de nombreux tests.

Pour cette phase, trois indices ont été calculé afin de discriminer les différents objets. Le NDVI (Normalized Difference Vegetation Index) est obtenu par la formule suivante : $NDVI = ([PIR] - [rouge]) / ([PIR] + [rouge])$. Il permet de caractériser les espaces couverts de végétation photosynthétiques (NDVI proche de 1) mais également les étendues d'eau lorsque le NDVI est proche de -1.

LE NDWI (Normalized Difference Water Index) est obtenu par la formule suivante : $NDWI = ([PIR] - [vert]) / ([PIR] + [vert])$. Il permet de caractériser la teneur en eau des végétaux et ainsi de différencier différents types de cultures. Le troisième paramètre est le MSAVI₂ (Modified Soils Adjusted Vegetation Index) est obtenu par la formule (Qi et al., 1994): $MSAVI_2 = (2 * [PIR] + 1 - ((2 * [PIR] + 1)^2 - 8 * ([PIR] - [rouge])))^{0.5}$. il permet de caractériser la végétation en tenant compte de l'humidité du sol.

Comme le présente l'arbre dichotomique de décision (en annexe 5 pour une question de place), les paramètres utilisés pour la discrimination des classes peuvent être basé sur les propriétés spectrales des objets (moyenne de bandes, NDVI...) ou des propriétés de textures (Homogénéité, entropie...). La définition de

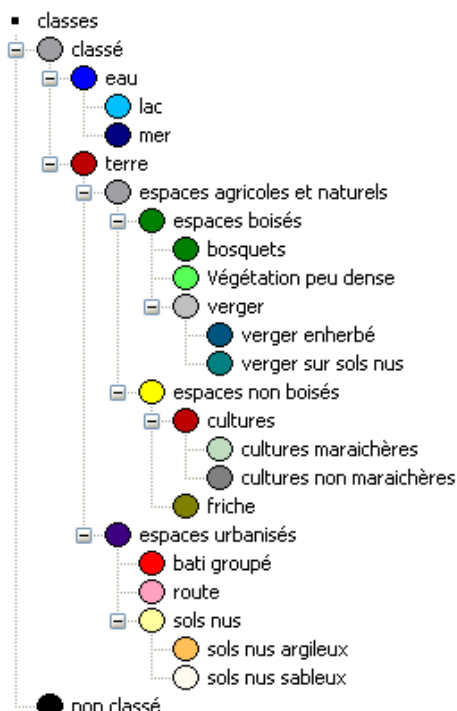


Figure 22 : Arbre de classification

chacun des paramètres utilisés dans cette classification est jointe au schéma ainsi que les valeurs seuils dans chaque cas.

Ces nombreuses tentatives ont permis de réaliser la carte d'occupation du sol présentée dans la figure 22.

Les classes utilisées sont plus nombreuses que les classes nécessaires à la suite de l'étude mais la sur classification de certains objets a permis une meilleur efficacité de la classification. En effet, les classes de bosquets, de végétation peu dense et de friche sont regroupées sous l'appellation 'Végétation naturelle' pour l'analyse.

Néanmoins, de par leur caractéristiques spectrométriques différentes, il a été très difficiles de les regrouper en une unique classe pour la classification. De même la nécessité d'obtenir des classes particulières (la mer par exemple) a entraîné la séparation de blocs de classification homogènes (comme l'eau) en sous classes. C'est le cas de la classe 'mer' et de la classe 'sols nus sableux'.

Une fois la classification réalisée, nous avons mis en place une matrice de confusion afin d'en mesurer la qualité.

Les classes ainsi définies ont permis la réalisation de la classification présentée par la figure 23.

La réalisation de la classification de l'image entière est une étape très importante pour l'acquisition des

paramètres environnementaux qui seront intégrés à l'étude statistique. Néanmoins, il est impensable d'y accorder entière confiance sans effectuer une vérification à partir de données vérifiées sur le terrain.

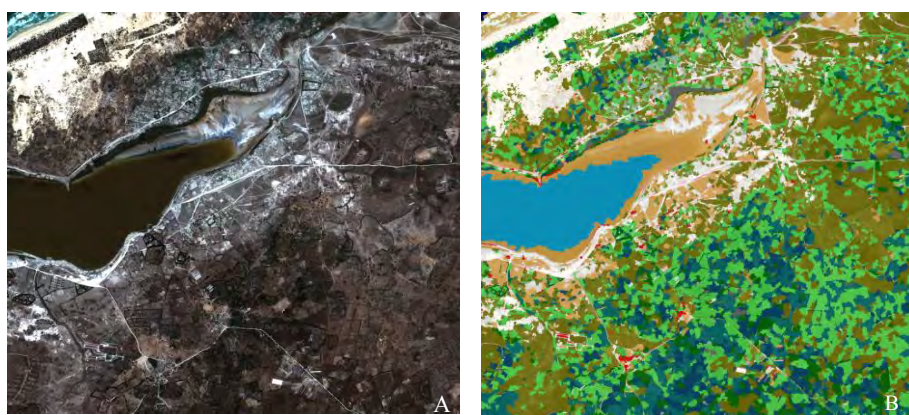


Figure 23 : Exemple de classification (A : image brute, et B : image classifiée)

3. Différents niveaux de classification

La classification a été détaillée au-delà des exigences de l'analyse statistique. Ceci nous permet de faire des regroupements selon différents critères. Ainsi, 4 niveaux de classification ont pu être testés.

Tableau 4 : Description des différents niveaux de classification

niveau 0	niveau 1	niveau 2	niveau 3
eau	eau	eau	Mer Lac
espaces non-boisés	espaces cultivés	cultures	cultures non maraichères cultures maraichères
espaces boisés		Verger	verger sur sols nus Verger enherbé
espaces non boisés	Végétation naturelle	Végétation naturelle	Friche basse
espaces non-boisés			Végétation peu dense
espaces boisés			Bosquets
espaces urbanisés	espaces urbanisés	espaces urbanisés	bati groupé Route
sols nus	sols nus	sols nus	sols nus argileux sols nus sableux

Le niveau 3 (voir tableau 4) correspond au niveau le plus détaillé et présente l'intégralité des couches définies dans la classification. Pour les trois autres niveaux, la mer et le lac ont été regroupés sous l'appellation 'eau', les bâti groupés et les routes sont été regroupés sous le nom d'espaces urbanisés' et les sols nus argileux et sableux ont été réunis sous le nom de 'sols nus'. Parmi les quatre classifications, seules

la classification de niveau 2 permet de distinguer les vergers et les cultures. C'est donc à partir de ce niveau que seront effectués la collecte des indicateurs environnementaux et la suite de l'analyse.

Une fois le niveau de classification fixé, il est important de collecter les données environnementales.

4. Collecte des données environnementales

La recherche de la vérité terrain permet de valider la classification et de connaître la part de segments (objets de segmentation) mal classés et quelles sont les classes qui sont le plus confondues. Pour réaliser cette vérité terrain, une série de sorties ont été effectuées ainsi que la création d'un fichier regroupant les points de références répartis sur l'ensemble de l'image.

1). Sorties terrain

Les sorties sur le terrain se sont faites après la dernière phase des observations d'insectes dans les parcelles. Cette série de déplacements a permis de repérer les différents ensembles paysagers existants et de faire le lien entre les structures visibles depuis le ciel et l'occupation du sol. Des planches papier de chacune des zones comprises dans le rayon d'un kilomètre autour de chaque parcelle ont été réalisées afin de nous assister dans cette tâche.

Avec ces planches, il a été possible de localiser et d'identifier l'ensemble des parcelles de crucifères (choux et navets) existant dans ces zones de 1 km. La localisation de chaque parcelle a été effectuée par la prise d'un point GPS à chaque coin de la parcelle. Ces points ont ensuite été reliés entre eux afin de fournir l'emprise exacte de chaque parcelle. L'acquisition des surfaces plantées en choux et en navets autour des parcelles d'intérêt a permis de calculer les indices relatifs aux plantes hôtes (voir partie IV.2.5 : Variables environnementales).

A la suite de ces sorties, il a donc été possible de réaliser une couche regroupant les points de référence correspondant à la réalité observée sur le terrain.

2). Ensemble des point de références

Comme le montre la figure 24, les points de références ont été placés de sorte à éviter au maximum les zones ayant servi à la mise au point de la classification. La limitation du nombre de points dans ces zones permet d'éviter les phénomènes de sous-estimation des zones mal classifiées.

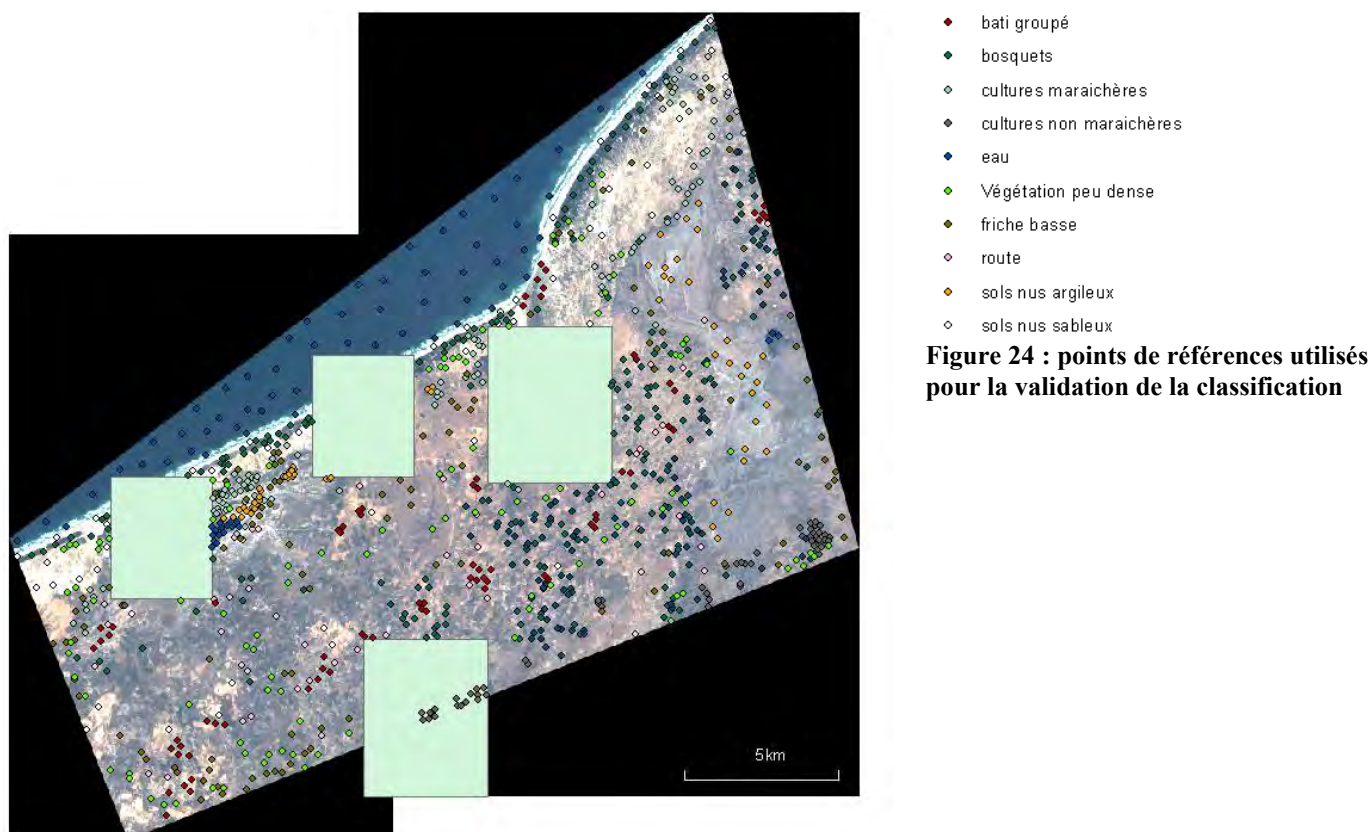


Tableau 5 : Nombre de points de référence alloués par classe

Classe	nb de points
bâti groupé	75
Bosquets	105
Cultures maraichères	105
Cultures non maraichères	60
Lac	24
Mer	51
Végétation peu dense	105
Friche basse	110
Route	50
sols nus argileux	60
sols nus sableux	50
verger enherbé	105
verger sur sols nus	100
Toutes les classes	1000

On peut observer que la répartition des points n'est pas homogène sur l'ensemble de l'image. Par exemple, les points de référence pour la classe 'Lac' sont regroupés dans un unique secteur. Ceci est dû au fait que le seul lac de l'image soit le Lac Rose et que celui-ci soit en partie contenu dans une zone de test. Des contraintes paysagères expliquent donc la répartition singulière de certaines classes (sols nus sableux au bord de la mer, sols argileux sur les zones de lacs asséchés, points du bâti groupé souvent rassemblés en amas...). Malgré ces contraintes, les points de références ont été répartis de manière la plus homogène possible, pour chaque classe.

De plus, une répartition du nombre de points selon la classe a été faite afin d'obtenir un total de points précis et une représentativité de chaque classe proportionnelle à la superficie qu'elle occupe dans

l'image. Le tableau 5 présente le nombre de points ainsi répartis selon la classe considérée. Ainsi, une classe recouvrant peu de surface (classe 'lac' ou 'route' par exemple) a reçu peu de points de référence en comparaison d'une classe plus représentée dans l'image (classe 'verger enherbé' ou 'friche basse').

Une fois les points de références positionnés, il est important de mettre au point une procédure de validation de la classification afin de s'assurer que la classification automatique retourne des résultats satisfaisants.

5. Validation de la classification

Un simple test permet de valider ou non une classification à partir de données de référence. Il s'agit du calcul de la matrice de confusion et des coefficients qui s'y rattachent.

Pour réaliser cette matrice de confusion, il a été nécessaire d'exporter la classification depuis e-cognition au format shapefile afin d'y accéder dans ArcGIS. Pour cela, il est

nécessaire d'exporter également le nom de la classe pour chaque segment. Une fois la couche de classification ouverte dans ArcGIS, il est possible d'y superposer les points de référence

Tableau 6 : Extrait de la table attributaire après jointure

classe référence	classification
bâti groupé	bâti groupé
bosquets	verger enherbé
bosquets	bosquets
cultures maraichères	cultures maraichères
friche basse	friche basse
mer	mer
route	sols nus sableux
sols nus argileux	sols nus argileux

et de réaliser une jointure spatiale selon l'emplacement. Ainsi, il est possible de voir pour chaque point de référence sa classe propre et la classe attribuée au polygone dans lequel il est contenu (voir tableau 6). On peut ainsi voir les premières confusions même si la vue d'ensemble est complexifiée par le nombre de points.

A partir de cette liste, il est possible de réaliser d'abord une matrice de distribution (voir tableau 7) grâce à la fonction 'tableau croisé

dynamique' d'Excel. Ce tableau présente pour chaque cellule le nombre de points appartenant à la classe de référence (en ligne) et à la classe donnée par la classification (en colonnes).

Tableau 7 : Matrice de distribution des points

Référence / classification	mer	bâti groupé	bosquets	cultures mar.	cultures non mar.	Végétation Peu dense	friche basse	lac	route	Sols argile	sols sable	verger herbe	verger nus	Total
mer	51													51
bâti groupé		26				1	8		1	24	15			75
bosquets			64	4	6	9	7					10	5	105
cultures mar.			1	46	33	12	2					1	10	105
cultures non mar.			2	3	32	3	6			12		2		60
friche arborée			7		1	29	26			3	3	18	18	105
friche basse						6	71			30		1	2	110
lac								6	1	13	4			24
route		6				1	6		13	17	5	2		50
sols argile							1			44	15			60
Sols sable							6				44			50
verger herbe			5			25	4					39	32	105
verger nus						30	28			3	3		36	100
Total	51	32	79	53	72	116	165	6	15	146	89	73	103	1000

Cette matrice de distribution a alors été intégrée dans un calculateur de matrice de confusion en ligne¹³. Ceci a permis de récupérer les deux indicateurs nécessaires à la validation de la classification (Landis and Koch, 1977). Le premier indice est le 'Overall accuracy' (Alberg et al., 2004). Cet indice d'exactitude globale représente le ratio (en pourcentage) entre le nombre de points correctement classés (valeurs en diagonal) et le nombre total de points. Plus la valeur est proche de 100%, plus la classification est exacte. Le second indice est le coefficient kappa. Cet indicateur représente le pourcentage de l'accord maximum corrigé de ce qu'il serait sous le simple effet du hasard.

Une fois la classification confrontée à la vérité terrain et validée, il est possible d'extraire les variables environnementales qui entreront dans la suite de l'analyse des données.

6. Choix des variables environnementales

Les variables environnementales intégrées à l'étude peuvent être classées en deux groupes, les indices caractérisant l'occupation des sols et ceux caractérisant le paysage.

1). Indices de composition du paysage

Cette catégorie regroupe cinq variables. La première est particulière puisqu'elle représente la distance en mètres entre la parcelle considérée et le segment classé en 'mer' le

¹³ <http://www.dicom.uninsubria.it/~marco.vanetti/cfmatrix/>

plus proche. Les autres variables représentent la surface occupée par une classe spécifique (ou un groupe de classes) sur une surface totale donnée. Ainsi, contrairement à la variable 'distance à la mer', les autres variables surfaciques sont comprises entre 0 et 1.

Ainsi, les proportions de plantes-hôtes (choux et navets), de verger (enherbé et sur sols nus), de cultures (maraichères) et de végétation naturelle (bosquets, friche basse et végétation peu dense) ont été calculées pour des surfaces correspondant aux zones tampons de 250, 500 et 1000 m de rayon autour des parcelles. Les indices caractérisant l'occupation des sols autour de chaque parcelles sont au nombre de 13 (trois pour chaque variables surfaciques et une pour la distance à la mer). Cette liste est complétée par une série d'indices caractérisant le paysage.

2). Indices de diversité

$$H' = - \sum_{i=1}^S p_i \log_2 p_i$$

Figure 25 : Formule de l'indice de Shannon

Cette catégorie regroupe 3 indices synthétiques qui caractérisent la configuration du paysage. Pour cette étude, 3 indices ont été choisis.

Le premier indice est l'indice de Shannon-Weaver (Spellerberg and Fedor, 2003). Comme le montre la figure 25, l'indice de Shannon (H') dépend de la proportion de chaque classe par rapport au nombre total de classes (p_i). Ce paramètre p_i est aussi appelé richesse spécifique. Cet indice permet de quantifier l'hétérogénéité de la biodiversité d'un milieu d'étude.

Le deuxième indice mis en œuvre est l'indice de fragmentation du paysage qui quantifie le morcellement du paysage. Il permet ainsi d'évaluer la différence qu'il existe entre le nombre de polygones après fusion (regroupement selon la classe s'il y a une limite commune) et le nombre de polygones totaux. En parallèle, l'indice de complexité quantifie la fragmentation du paysage mais rattaché au nombre de classes présentes. Ainsi, un paysage très fragmenté mais ne montrant que 3 classes aura un indice de complexité plus faible qu'un paysage également fragmenté avec 10 classes. Les indices ont été calculés pour les zones tampons de rayon 250, 500 et 1000 m

Une fois le traitement de l'image et le calcul des indicateurs environnementaux effectués, il est nécessaire de voir quels sont les protocoles mis en place pour la gestion et l'analyse des différentes variables.

C. Analyse des données

L'objectif principal du stage est de déterminer les liens existant entre les pratiques culturelles, le paysage et la régulation des ravageurs. Pour y arriver l'analyse des données est nécessaire. Mais avant d'analyser les données, il a été nécessaire de les collecter, de les traiter et de les organiser. Ici, l'intégration des données s'est faite sous Access et l'analyse des données s'est fait sous R.

1. Compilation sous Access

Avant l'analyse statistique des données, trois étapes ont été nécessaires, l'intégration des données de terrain et de laboratoire sous Access, le choix et le calcul des variables à expliquer, et enfin la mise en forme des données pour l'analyse statique que nous avons effectués avec le logiciels R.

1). Gestion et relations des données

L'unité de base de l'étude est la parcelle. Ce point est bien visible lors de l'observation des relations liant les différentes tables sous Access (voir figure 26). Pour toutes les tables, la

clé primaire se compose ou contient le l'identifiant de la parcelle.

En plus de ce numéro de parcelle, la table principale 'Parcelle' contient toutes les données relatives aux pratiques culturales.

Grace à la présence du numéro de parcelle dans chaque table, il est possible de faire le lien entre les différentes tables et de mettre en place différents indicateurs et de comparer ensuite les parcelles entre elles à la lumière de ces indicateurs.



Figure 26 : Relations entre les tables de la base de données Access

2). Indicateurs biologiques

La mise en relation des différentes informations recueillies sur le terrain et au laboratoire a permis de calculer des indicateurs synthétiques relatifs au ravageur et au parasitisme. Il a ainsi pu être calculé deux variables d'intérêt afin de développer la suite de l'analyse.

La première variable est le taux de plant infesté. Il s'agit du nombre de plants observés présentant au moins une larve de l'espèce *Plutella xylostella* ramené au nombre de plant observés au total. Cette variable est calculée pour chaque parcelle sur l'ensemble des trois observations. La deuxième variable calculée à partir des données collectées est le taux de parasitisme. Il s'exprime par le nombre d'individus ayant été parasité sur le taux d'individus collectés par parcelle.

Une fois les indicateurs biologiques évalués et compilés, il est important de voir sous quelle forme les mettre en afin de faciliter l'analyse sous R

3). Mise au propre des données

Le logiciel R accepte les formats texte et csv où les données sont présentées sous forme de tableaux. Le format texte présente l'avantage d'être peu gourmand en termes de volume mais présente également le défaut d'être moins lisible sans conversion vers Excel. Le format csv est quant à lui programmé pour s'ouvrir automatiquement sous Excel.

Les deux formats sont compatibles de manière équivalente avec le logiciel R et le choix du format de données est plus une question d'habitude de la part de l'utilisateur que d'un choix de performances.

Néanmoins, des caractéristiques particulières sont imposées aux données afin de permettre une lecture sans défaut par le logiciel R. Premièrement, les noms de variables longs sont un frein au bon déroulement des programmes car souvent sources d'erreur. C'est pourquoi il a été décidé d'attribuer à chaque variables un code alphanumérique (voir annexe 6 : Glossaire des variables utilisées pour l'étude statistique). Ainsi, les variables biologiques à expliquer reçoivent un code commençant par la lettre V, les variables culturelles un code avec la lettre P (comme 'Pratiques') et les variables environnementales un code avec la lettre E.

Ensuite, le format des nombres réels est réglementé par le logiciel R. En effet, dans certains cas, la virgule sert de séparateur de colonnes. Il peut donc avoir une confusion entre la virgule utilisée comme séparateur des colonnes et la virgule utilisée comme séparateur des unités et des décimales dans un nombre. C'est pour cela que les fichiers exportés en vue d'être analysé par R ont été transformé afin de changer la virgule servant de séparateur entre l'unité et les décimales en points ('16,5' devient '16.5').

Enfin, il est important d'éliminer les lignes et colonnes totalement vides en apparence (espaces invisibles en particulier) afin de limiter le nombre de cellules apparaissant comme non renseignées (NA) lors du transfert vers R. En effet, beaucoup de fonctions de R ne reconnaissent pas ces données et ne fonctionnent donc pas en leur présence.

Une fois l'intégralité de ces points vérifiés, il a été aisé de créer les fichiers présentés en annexe 7, 8 et 9, respectivement l'ensemble des données relatives aux variables à expliquer, aux variables environnementales et aux variables culturelles.

Ainsi, l'analyse des données a pu être effectuée sans avoir à remanier les données pour chaque bloc de codes.

2. Analyse statistique sous R

Les modèles statistiques utilisés ici sont de la famille de régression linéaire multiple. En effet, ce type de modèles est appliqué dans le cas où les variables à expliquer sont de nature continue et les variables explicatives de type continues ou ordonnées, ce qui est le cas ici.

L'algorithme complet n'est pas présenté ici puisque sa taille le rend particulièrement illisible. Ainsi, pour une meilleure compréhension de l'algorithme, seules les lignes de commandes permettant l'appel des données et des librairies, l'étude des variables, la recherche des modèles linéaires (selon la loi gaussienne) et la recherche des modèles selon la loi binomiale sont présentées.

1). Appel des données et librairies

```
#Appel des librairies
Library (nom-librairie)

# Chemin d'accès et nom du fichier source
setwd (dir = "chemin-d'accès")
db<read.csv2('données.csv',header=T)

#visualisation des données
Summary(db)

#attribution d'un nom aux variables
Variable<-db$Variable
```

Figure 27 : Code R pour l'appel des données et des librairies

La première étape de toute analyse statistique est le chargement des librairies contenant les fonctions utilisées et l'appel des données. Même si le chargement des librairies est dépendant des fonctions appelées, il est important d'en faire la liste et de connaître les fonctions qui s'y rattachent. L'ensemble des lignes de commandes utilisées dans cette partie sont présentées dans la figure 27 par les lignes de code bleues.

Six librairies ont ainsi été chargées afin de mener à bien l'étude statistique. La première librairie nécessaire pour le traitement des

données est la librairie 'RODBC' (R Open Database Connectivity) qui permet la lecture

des différents types de données. Ensuite, il est nécessaire de charger un lot de librairies permettant l'étude statistique proprement dite avec la librairie 'lme4' qui permet de réaliser les analyses statistiques relatives aux modèles linéaires, la librairie 'MASS' (modern applied statistics with S) et la librairie 'MuMIn' (multi model inference) pour les modèles plus complexes. Enfin, il a été nécessaire de faire appel à la librairie 'Hmisc' pour la visualisation des graphiques et à la librairie 'aod' pour les tests de correspondance entre la prédiction extraite du modèle testé et les valeurs réelles.

Une fois les librairies connues et chargées dans l'environnement de travail de R, il est important de renseigner le chemin d'accès des données par la fonction `stwd` (ce qui évite de devoir modifier manuellement le dossier de travail à chacune des ouvertures du logiciel. Après avoir défini l'espace de travail, il est nécessaire d'appeler le fichier de données avec la `read.csv2` (dans le cas des fichiers .csv) ou `read.table` (pour les fichiers .txt).

Les deux dernières étapes de l'appel des librairies et des données sont la visualisation des données (fonction `summary`) et l'attribution d'un nom court pour chacune des variables. Ces deux étapes permettent de vérifier si le type des données est correct (type numérique ou facteur) et de simplifier la suite des codages.

Une fois les données et toutes les fonctions nécessaires intégrées au logiciel, il est temps d'évaluer la corrélation entre les variables qui vont être intégrés dans l'étude.

2). Etude des corrélations entre les variables

L'étude du lien entre les différentes variables permet de visualiser les variables grandement corrélées et d'éliminer les variables redondantes.

```
# Correlation des variables
db2<-data.frame("var1"=var1, "var2"=var2,...,
stringsAsFactors=FALSE)

# Tracé des relations entre les variables quantitatives
pairs(db2,pch = 21, bg = c("black"))
TableCor2<-rcorr(as.matrix(db2))
write.matrix(TableCor2, file = "correlation", sep = "\t")
```

L'étude des corrélations entre les variables nécessite la mise en place d'un cadre de lecture strict afin de permettre la lecture des données et le bon fonctionnement des fonctions (fonctions `data.frame`).

Une fois cette base mise en place, il est possible de tracer le graphique représentant l'ensemble des tests de corrélation pour tous les couples de variables possibles (fonction 'pairs'). La fonction 'rcorr' permet de

l'ensemble des modèles sont exporté vers un fichier Excel afin de pouvoir sélectionner manuellement les meilleurs modèles.

```
# Calcul des modèles
lm1 <- lm(formula = Var ~ ., data=db2)
ms1 <- dredge(lm1, trace=TRUE, rank = "AICc")

ms1 <- dredge(lm1, extra = list(
  "R^2", "*" = function(x) {s <- summary(x)
  c(Rsq = s$r.squared, adjRsq = s$adj.r.squared,
  F = s$fstatistic[[1]])}))

write.matrix(ms1, file = "AICc.xls", sep = "\t")

# Calcul des valeurs prédites par le modèle
fm01<- lm(formula = V1 ~
  formule du modèle, data=db2)
Infest_predite01 <- predict(fm01)

cor(V1,Infest_predite01)
exp(cbind(OR = coef(fm01), confint(fm01)))

##Graphique des valeurs prédites
par(mfrow=c(2,2))
plot(fm01)
plot(predict(fm01),V1, main="Nuage de points ")
abline(lm(V1~Infest_predite01), col=2)
```

Figure 29 : Code R pour la recherche des modèles selon la loi gaussienne

4). Recherche de modèles suivant la loi binomiale

```
# Calcul des modèles
lm1 <- glm(formula = cbind(V11,V12) ~ .,
  family=binomial, data=db2)

# Calcul des valeurs prédites par le modèle
fm01<- glm(formula = cbind(V11,V12) ~
  E10+E21+E5+E7+P7, family=binomial, data=db2)
```

Figure 30: Codes R pour la recherche de modèles selon la loi binomiale

de recherche de modèles. En effet, la fonction 'lm' est remplacée par la fonction 'glm'. Cette fonction permet de rechercher des modèles selon différentes lois statistiques. Il est possible d'appliquer la loi gaussienne (comme dans la fonction 'lm') mais aussi la loi de Poisson et la loi Binomiale. Cette dernière loi est utilisée ici, c'est pourquoi le paramètre 'binomial' est intégré dans la partie 'family' de la formule.

De plus, une modification de la forme des variables à expliquer a été nécessaire afin de réaliser cette analyse. En effet, la fonction binomiale demande comme variables d'entrée les valeurs de réussite et échec du test auquel la variable se réfère. Ainsi, pour le taux de plant infesté V1, deux variables supplémentaires ont été créées, V11 pour le nombre plant présentant au moins une larve, et V12 le nombre de plant ne présentant pas de larve. Une fois ces deux paramètres intégrées comme variables d'entrée à la place de V1, il est possible de réaliser la recherche de modèle, le calcul des valeurs prédites et leur affichage graphique

Le second groupe de codes permet le calcul de valeurs prédictives selon le modèles entré dans la partie 'formule du modèle'. Ce volet de codes permet ainsi de calculer la corrélation existant entre les prédictions issues de ce modèle et les valeurs de la variable obtenues par l'expérimentation.

Enfin, le dernier groupe de code regroupe l'ensemble des commandes liées au tracé des graphiques de corrélation grâce aux fonctions 'plot' et 'abline'

Une fois cette étape réalisée, il a été évident que la prédiction des variables grâce aux modèles linéaires gaussiens n'était pas la méthode optimale (voir partie résultats). Il a alors été décidé de réaliser la même étude avec un algorithme permettant la recherche de modèles suivant une loi binomiale.

La série d'algorithme permettant la recherche des modèles peut être adapté afin de rechercher les modèles les plus fiables suivant une loi binomiale. La figure 30 présente donc uniquement les codes ayant été modifiés entre les deux versions.

La première modification à apporter est le changement de la fonction

comme dans le cas des modèles suivants une loi gaussienne. Les variables environnementales et culturelles intégrées dans la recherche des modèles selon la loi binomiale sont les mêmes que celles utilisées pour la recherche des modèles de la loi normale.

Une fois l'ensemble de la méthodologie étudiée, il est temps de visualiser et d'interpréter les résultats obtenus lors de cette étude.

V. Analyse des résultats

L'analyse des résultats suit le même plan que l'étude de la démarche méthodologique et le déroulement du stage. Ainsi, il est plus aisé de suivre le cheminement de l'analyse et de comprendre les résultats obtenus. Il sera donc abordé d'abord les résultats relatifs aux sorties sur le terrain et au laboratoire, ensuite les résultats concernant le traitement de l'image et enfin les résultats obtenus à partir de l'étude statistique.

A. Résultats issus des données de terrain et de laboratoire

Les premiers résultats disponibles dans cette étude sont les résultats relatifs aux sorties sur le terrain et au laboratoire. Il est possible de les classer en trois groupes selon le thème qu'ils abordent. En premier viennent les résultats se rapportant aux parcelles elles-mêmes, ensuite, les résultats relatifs aux individus étudiés et enfin les résultats concernant les pratiques culturelles.

1. Parcelles

Les parcelles sont des éléments riches en enseignement une fois que les données qui s'y rattachent ont été analysées. Les premiers résultats obtenus sont en relation avec les caractéristiques de la parcelle elle-même alors que les seconds sont plutôt tournés vers la présence et la répartition des ravageurs entre les parcelles.

1). Caractérisation des parcelles

Le premier jeu de résultats s'intéresse donc aux caractéristiques des parcelles. Les 31 parcelles ont été choisies et réparties sur la zone d'étude selon les facteurs détaillés dans la partie méthodologie (choux jeunes, à plus de 1 km d'une autre parcelle et données culturelles accessibles). Ainsi, la répartition des parcelles a été fixée selon le plan représenté par la figure 31. Pour bien comprendre la problématique liée au périmètre de 1000m autour de chaque parcelle, les zones tampons correspondant à cette distance ont également été représentées sur la carte grâce à un trait rouge. On observe quelques zones où des chevauchements entre zones tampons existent. Même si ce chevauchement a été un facteur discriminant dans le choix des parcelles, la nécessité fait souvent loi et ces chevauchements ont été tolérés quoiqu'ils soient limités au maximum.

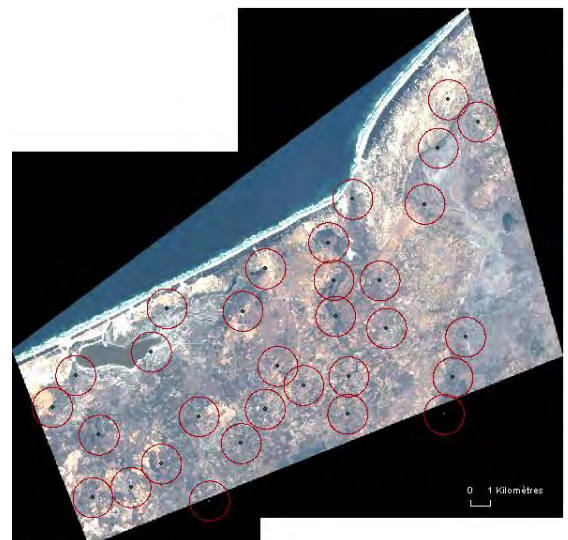


Figure 31 : Répartition spatiale des parcelles observées

Taille de la
pomme (cm)

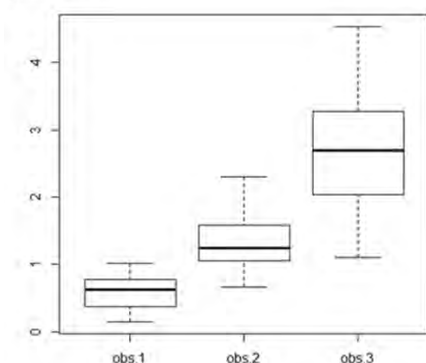


Figure 32 : Evolution de la taille des choux selon le numéro d'observation

Nombre moyen de larves de
Plutella xylostella par plant

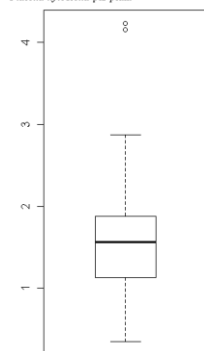


Figure 33 : Répartition du nombre moyen de larves par plant

Le second résultat exploitable concernant les parcelles est la taille des choux. En effet, cette mesure est effectuée de manière systématique et donc disponible pour chaque échantillon (2730 mesures en tout). Il est alors possible d'effectuer des statistiques sur l'évolution de la taille de la pomme du chou (partie consommable) en fonction du temps. Comme le montre la figure 32, cette notion de temps est représentée par le numéro de l'observation puisque pour certaines parcelles, l'observation 1 s'est faite en même temps que l'observation 2 d'autres parcelles. Cet indicateur temporel permet donc de régulariser ces différences. On peut ainsi observer une augmentation hyperbolique de la taille de la pomme en fonction du temps. Cette progression représente le début de la phase de développement végétatif du chou, qui est la

période la plus critique en matière de pression des ravageurs. La temporalité de l'étude est donc en adéquation avec les conditions biologiques qui lui sont nécessaires.

Une fois les choux observés et échantillonnés, il est possible de connaître le nombre d'insectes présents sur chaque chou et ainsi d'accéder au nombre moyen de ravageurs (ici uniquement les individus de l'espèce *Plutella xylostella* par plant à l'échelle de la parcelle. La figure 33 présente cette répartition et permet de constater une variabilité intéressante de ce paramètre. En effet, on peut voir des parcelles présentant beaucoup de larves par plant (points au-dessus du neuvième décile) et d'autres proches de l'absence de ravageurs. Ce constat est encourageant puisqu'il laisse présager une variation des taux d'infestation des parcelles assez importante.

2). Caractérisation de l'infestation

Une fois les paramètres des parcelles étudiées, il est intéressant de se pencher sur la notion d'infestation et de voir quels sont les résultats qui en émergent.

Le premier résultat observable en termes d'infestation est bien sûr la répartition de ce taux par parcelle. La figure 34 présente ce taux. On peut observer une forte variation de cet indicateur (de 0.3 à 0.9). Le taux d'infestation étant l'une des variables intégrées à la recherche de modèle, une variabilité importante permet d'espérer des résultats plutôt intéressants lors de l'analyse.

De plus l'absence de valeurs exceptionnelles (au-dessus ou en dessous de la boîte de Tukey) limite le risque de trouver des valeurs pivots dans l'analyse statistique et limite le risques de trouver des modèles déséquilibrés par ces valeurs.

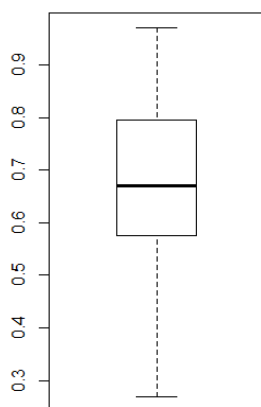


Figure 34 : Répartition du taux d'infestation par parcelle

Une fois la répartition statistique des taux d'infestation connue et analysée, il est intéressant d'observer comment cette répartition se comporte dans l'espace. Pour cela, la carte présentée dans la figure 35 a été conçue afin de visualiser la répartition spatiale des taux d'infestation à travers le code couleur. Ainsi, la couleur rouge devient de plus en plus foncée à mesure que le taux d'infestation augmente.

On peut observer que les points dont le taux d'infestation est le plus faible sont regroupés sur le bord d'une zone sombre. Il s'agit de la bordure d'un lac asséché dont les sols sont majoritairement argileux. On peut également observer le regroupement des valeurs les plus fortes (au-dessus de 0.70) le long d'un axe transversal allant du coin Sud-ouest à la mer vers le milieu de la carte. Sur le terrain, il est possible de voir que l'intégralité des parcelles entrant dans ce cas de figure est proche d'une route goudronnée.

Une fois ces premières constatations faites, il est de plus en plus intéressant de réaliser l'analyse statistique des variables environnementales afin de savoir quels sont les paramètres qui influencent réellement le taux d'infestation. De plus, il est important de comprendre les résultats émanant du laboratoire en termes d'individus collectés et de parasitisme.

2. Individus

Le deuxième type de résultats obtenus en analysant les données provenant du terrain concernent les individus qui ont été collectés à travers leur caractéristiques et celles des parasitoïdes qui les accompagnent.

1). Caractérisation des individus collectés

La première caractéristique visible des individus observés est l'espèce. La répartition présentée dans la figure 36 a été obtenue à partir des relevés de populations effectués sur le terrain où apparaissent le nombre d'individus des espèces *Plutella xylostella*, *Hellula undalis* et *Chrysodeixis chalcites*. On peut ainsi observer que les individus de *Plutella xylostella* sont représentés en majorité (98.89%).

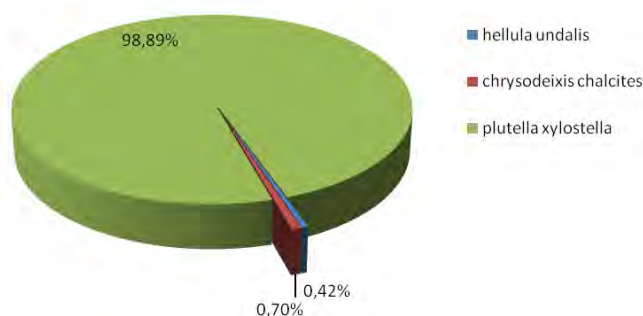


Figure 36 : Répartition des espèces observées

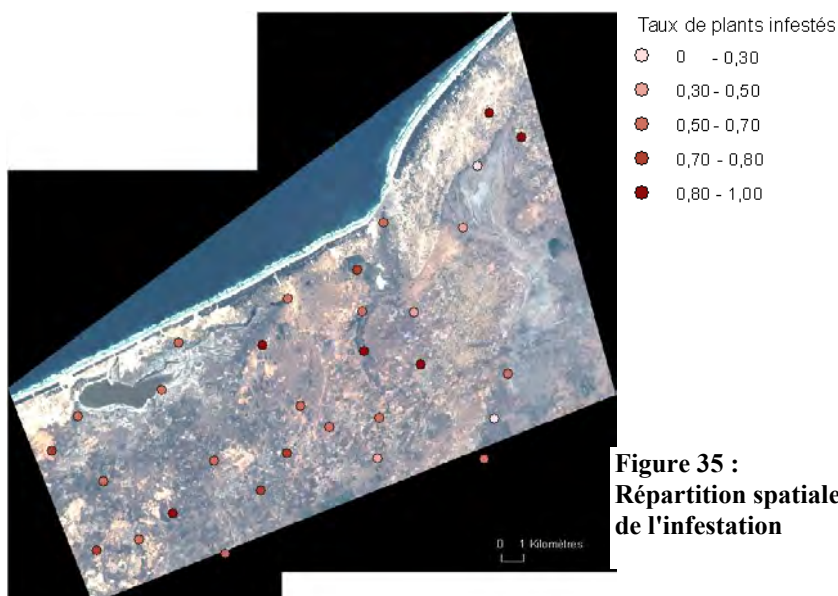


Figure 35 : Répartition spatiale de l'infestation

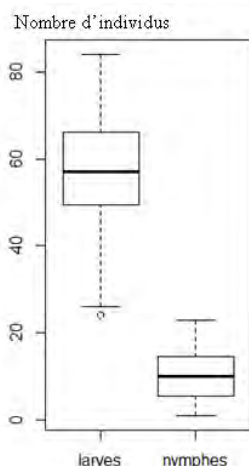


Figure 37 : Répartition des individus observés selon le stade de collecte

Néanmoins, ce chiffre de presque 99% regroupe les larves de *Plutella xylostella* mais également ses nymphes. Il est donc important de connaître la répartition de chacune de ces populations dans l'échantillon rapporté au laboratoire. La figure 37 présente ces résultats. A part une valeur exceptionnelle, il est évident que le nombre de larves ramenées au laboratoire est très supérieur au nombre de nymphes. Ceci est expliqué par le fait que les choux jeunes présentent significativement plus de larves de petites tailles que de nymphes. Ce rapport s'équilibre sur des plants plus âgés où les insectes ravageurs sont plus avancés dans leur cycle de développement.

La caractérisation du parasitisme doit se faire séparément sur ces deux groupes puis sur l'ensemble des individus afin de ne pas risquer de perdre de l'information importante.

2). Caractérisation du parasitisme

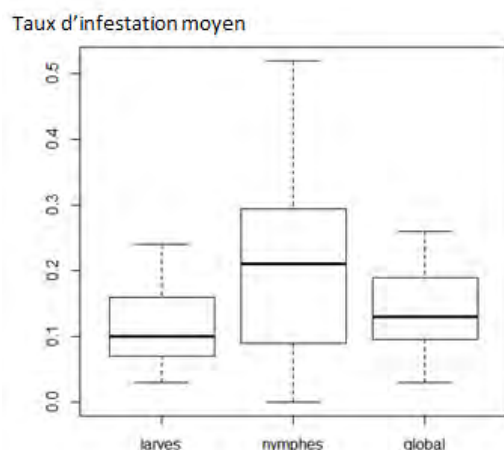


Figure 38 : Infestation parasitaire en fonction du stade du ravageur à la collecte

Même si le nombre d'individus parasités est faible face au nombre d'individus atteignant l'âge adulte (452 contre 1033, soit 18% contre 33% de la population totale), il est intéressant de caractériser ce parasitisme en termes de taux d'infestation parasitaire et de caractéristiques des parasitoïdes.

La figure 38 montre que le taux d'infestation parasitaire est différent selon que l'on considère l'ensemble des individus (global), le groupe des nymphes ou celui des larves. On observe que la réponse de la population globale est proche de celle des larves. Ceci est dû à la grande différence qu'il existe en termes d'effectifs entre les larves et les nymphes. Pour la suite de l'étude, aucune différence ne sera faite entre ces deux groupes et les variables seront calculer à partir de la population totale.

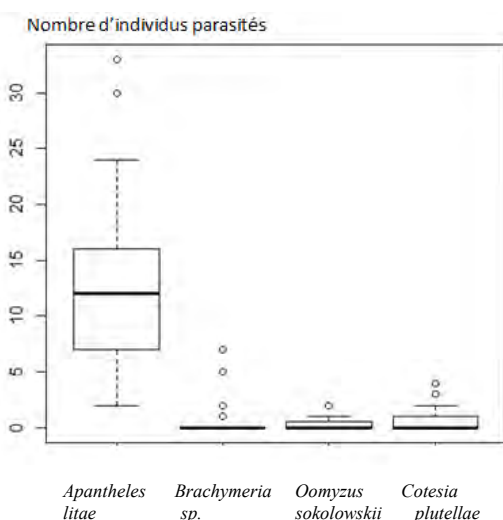


Figure 39 : Type de parasites impliqués dans l'infestation

Une fois la population parasitée connue, il est important de s'intéresser à la population des parasites. En effet, les quatre espèces de parasitoïdes présentent un nombre d'individus infestés différents. La figure 39 montre le nombre d'individus de l'espèce *Plutella xylostella* ayant été infestés par un parasitoïde en fonction de l'espèce de celui-ci. Ce choix d'axe est dicté par le comportement grégaire d'*Oomyzus sokolowskii* (plusieurs individus de parasitoïdes peuvent provenir de la larve d'un individu de *Plutella xylostella*). On peut observer que le parasitoïde *Apantheles litae* est impliqué dans un grand nombre d'infestation. Néanmoins, il est peu évident d'en tirer des

conclusions sûres et fiables de par le nombre de valeurs exceptionnelles présentes pour chaque modalité. L'approfondissement de l'étude avec l'obtention d'un échantillon plus large est préférable pour conclure sur cette partie.

Une fois les parcelles et les individus étudiés, il reste un dernier jeu de données issus des sorties de terrain à analyser. Il s'agit des données concernant les pratiques agricoles.

3. Pratiques agricoles

L'étude des pratiques agricoles permet de mettre en avant les premières conclusions concernant ces données. Ces conclusions seront complétées et approfondies au cours de l'étude statistique. Pour ces premières constatations, il est intéressant d'analyser le contexte culturel de la parcelle et ensuite de faire un point plus précis sur les produits phytosanitaires utilisés.

3). Caractérisation du contexte culturel

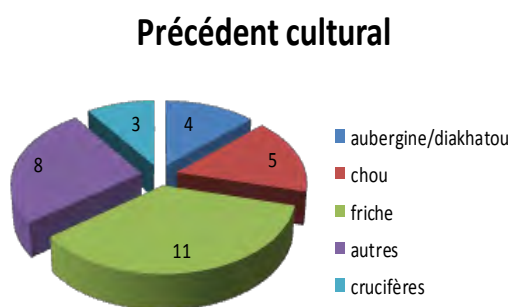


Figure 40 : Précedent culturel des différentes parcelles



Figure 41 : Type de sol des différentes parcelles

Le contexte culturel d'une parcelle influe de manière directe sur la plantation en cours. C'est pour cela que la connaissance de ce contexte est important pour la compréhension de l'état de la culture et par ainsi en partie de son infestation. Même si de nombreux paramètres impactent ce contexte culturel, deux facteurs principaux ont été intégrés à l'étude. Il s'agit du précédent culturel et du type de sol. Il est donc important d'analyser les résultats issus de ces deux paramètres.

Le précédent culturel de la plantation a un impact sur celle-ci puisqu'il conditionne les éléments nutritifs potentiellement présents dans le sol ainsi que leur quantité. Il a été démontré pour les céréales qu'une monoculture entraînait une dégradation de la qualité du sol et une diminution drastique des rendements obtenus. Dans le cas de notre étude (voir figure 40), une majorité des parcelles de choux sont cultivées avec un précédent de friche (absence de vraies cultures).

Les sols sont donc de composition variable sans carence particulière à craindre.

Néanmoins, les 8 parcelles ayant un précédent de crucifères ou de choux peuvent présenter plus de difficultés en raison de l'effet de monoculture. Ce paramètre sera à surveillé au moment de l'analyse statistique des données.

Le second paramètre influençant le contexte culturel est le type de sol. Ne pouvant entrer dans le détail fin du type de sol, les trois grandes catégories ont été utilisées, les sols argileux, les sols sableux et les sols sablo-argileux. On peut observer dans la figure 41 que la majorité des parcelles sont constituées d'un sol sableux. Il sera donc peu probable que ce paramètre sera utilisé pour discriminer les variables biologiques.

Un autre type de paramètres cultureux peut influencer les variables biologiques de l'étude. Il s'agit des traitements phytosanitaires.

4). Etude des produits insecticides

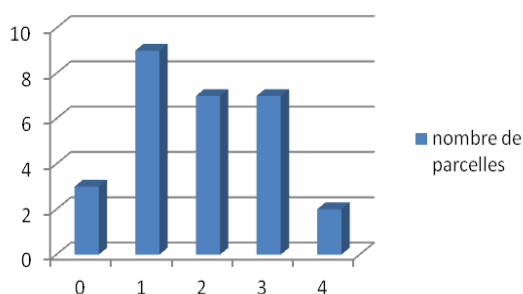


Figure 42 : Nombre de traitements phytosanitaires réalisés après le repiquage par parcelle

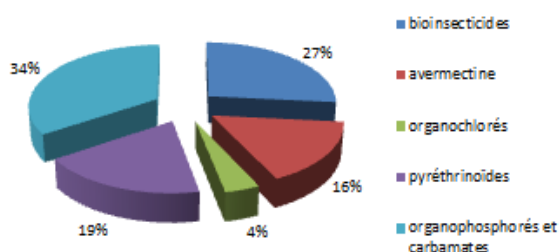


Figure 43 : Type de produits phytosanitaires utilisés

Les produits insecticides sont les outils les plus couramment utilisés dans la lutte contre les ravageurs. Il est donc important de connaître le nombre de traitements effectués sur une parcelle mais aussi quels sont les produits utilisés.

L'étude se faisant durant la période suivant le repiquage, il est important de comptabiliser le nombre de traitements insecticides que chaque parcelle a reçu durant cette période. La figure 42

nous montre que le nombre de traitements varie entre 0 et 4 sur la parcelle étudiée avec une répartition proche d'une courbe gaussienne. Il est donc probable que ce paramètre ait une influence sur les variables biologiques. Ceci est étayé également par les connaissances théoriques acquises sur le sujet.

Les différents produits insecticides recensés ont été regroupés en 5 classes selon la famille de leur matière active. Comme le montre la figure 43, deux types de produits sont majoritairement utilisés, les bio insecticides (comme le Bt ou les extraits de neem) et les organophosphorés et

carbamates (comme le diméthoate ou le métaphos). Même si les bio insecticides sont a priori plutôt sélectifs, donc peu toxiques pour les ennemis naturels, les organophosphorés sont des produits à large spectre. Ainsi, le paramètre 'sélectivité' des produits peut être intéressant à étudier plutôt que le nombre de traitements réalisés.

Une fois les données de terrain et de laboratoire étudiées et évaluées, il est temps de se pencher sur les résultats émanant du traitement de l'image satellite.

B. Résultats du traitement de l'image satellite

Les principaux résultats relatifs aux traitements de l'image satellite concernent la fiabilité de la classification. En effet, la suite de l'analyse des données environnementales repose sur le processus de classification et sa corrélation avec la vérité terrain. Pour appréhender cette fiabilité, différents niveaux de classification ont été mis en place. Il est ensuite important de comprendre les résultats obtenus par les matrices de confusion de ces différents niveaux de classification et d'expliquer les causes d'erreur les plus probables avant de pouvoir visualiser la classification finale.

1. Matrice de confusion

L'évaluation de la classification a été réalisée sur les quatre niveaux de segmentation en se basant sur un tableau à deux dimensions appelé matrice de confusion (tableau 8) qui met en relation la classification créée et la vérité terrain. La précision globale de validation (nombre de pixels bien classés/nb total d'individus) a été calculée sur chaque niveau de segmentation. Cependant, cet indice ne tient pas compte des erreurs d'omission et de commission. Le coefficient Kappa, quant à lui, prend en compte de ces erreurs en ligne et en

colonne : il permet d'exprimer la réduction proportionnelle de l'erreur obtenue par une classification, comparée à l'erreur obtenue par une classification complètement au hasard.

En effet, ils permettent de voir et de comprendre quelles sont les erreurs commises et quelle est leur impact sur la fiabilité de la classification.

Tableau 8 : Matrice de confusion pour la classification de niveau 2

	cultures	eau	Végétation naturelle	espaces urbanisés	sols nus	vergers	classification overall	producer accuracy
Cultures	114	0	26	0	12	13	165	69,09%
Eau	0	57	0	1	17	0	75	76,00%
Végétation naturelle	11	0	219	0	36	54	320	68,44%
espaces urbanisés	0	0	16	46	61	2	125	36,80%
sols nus	0	0	7	0	103	0	110	93,63%
Vergers	0	0	92	0	6	107	205	52,19%
truth overall	125	57	360	47	235	176	1000	
user accuracy	91,20%	100,00%	60,83%	97,87%	43,83%	60,79%		

Le tableau 8 présente la matrice de confusion relative au niveau de classification 2. On peut observer des corrélations plutôt élevées pour les classes ayant le plus d'intérêt pour l'étude (végétation naturelle, cultures et vergers). De plus, les confusions sont nettement visibles entre certaines classes. L'étude des causes de ces confusions se fera dans la partie suivante. Il est ici important de comprendre et d'analyser les deux indicateurs permettant de valider la classification. Ces indicateurs sont l'Overall Accuracy (exactitude relative) et l'indicateur Kappa.

Tableau 9 : Indices de classification

niveau	OA (%)	kappa
0	76.59	0.67
1	65.9	0.538
2	64.6	0.553
3	62.2	0.58

Le tableau 9 présente la valeur de ces deux indicateurs pour chacun des quatre niveaux de classification. On peut observer que l'exactitude relative diminue en même temps que le nombre de classes augmente. Ceci est dû à l'augmentation du nombre de confusion possible. L'évolution du kappa quant à lui est due aux différences existant entre la vérité terrain et ce qui a été trouvé par la classification.

La classification la plus à même de représenter les données à étudier est celle correspondant au niveau 2 (voir partie méthodologie). Néanmoins, le kappa à 0.55 indique un accord modéré. Faute de temps, la classification proposée a été validée en l'état mais il serait nécessaire d'affiner la classification afin d'améliorer les indicateurs (kappa au minimum à 0.61 pour un accord fort).

2. Causes d'erreur de la classification

Ici, seules les confusions impactant la classification au niveau 2 seront étudiées. Néanmoins les conclusions tirées ici sont applicables aux autres niveaux de classifications.

Trois grandes confusions sont visibles dans la matrice (plus de 30 individus mal classés). La première confusion impacte directement la classification puisqu'elle concerne les classes 'Végétation naturelle' et 'verger'. En effet, une confusion entre ces deux classes peut entraîner des erreurs dans le calcul des taux de vergers et de végétation naturelle. Cette confusion est due à la nature même des classes puisque les deux regroupent des espaces boisés denses. La deuxième confusion a également un fort impact sur l'analyse mais à moindre échelle puisqu'elle concerne la 'végétation naturelle' et le 'sol nus'. Cette confusion est due à la présence dans la classe de végétation naturelle de friches basses et donc non boisées. La dernière confusion n'a aucun impact sur la suite de l'analyse puisqu'elle concerne la classe des 'espaces urbanisés' et celle des 'sols nus'. Cette confusion est sans doute due à des effets de réflectance sur des toits ou des sols clairs.

De plus, il a été possible de comparer les zones de maraichages obtenues au cours des sorties sur le terrain avec les zones de maraichages obtenus par la classification. Une grande différence est visible entre ces deux jeux de données. Ceci peut être dû au temps de rotation des cultures maraichères. En effet, l'image a été acquise le 16 février et les données terrain autour du 15 mai. Dans ce laps de temps, il est possible que des zones jusque-là inexploitées soient converties en zones maraichères ou inversement, des zones maraichères abandonnées ou plantées d'espèces non recensées pendant les sorties. Les différences de localisation des zones maraichères sont donc explicables par ce fait.

Grâce à la connaissance des facteurs d'erreurs dans la classification, il sera possible de l'améliorer. Néanmoins, le choix a été fait de garder la classification issue du deuxième niveau de regroupement en l'état par souci de temps.

3. Classification finale

La classification sélectionnée pour la suite de l'étude est donc composée de 5 classes. Par souci de place, une version réduite de la figure représentant l'ensemble de la zone d'étude classifiée est présentée ici (figure 44). Une image en plus grand format a été portée en annexe 10 pour plus de lisibilité.

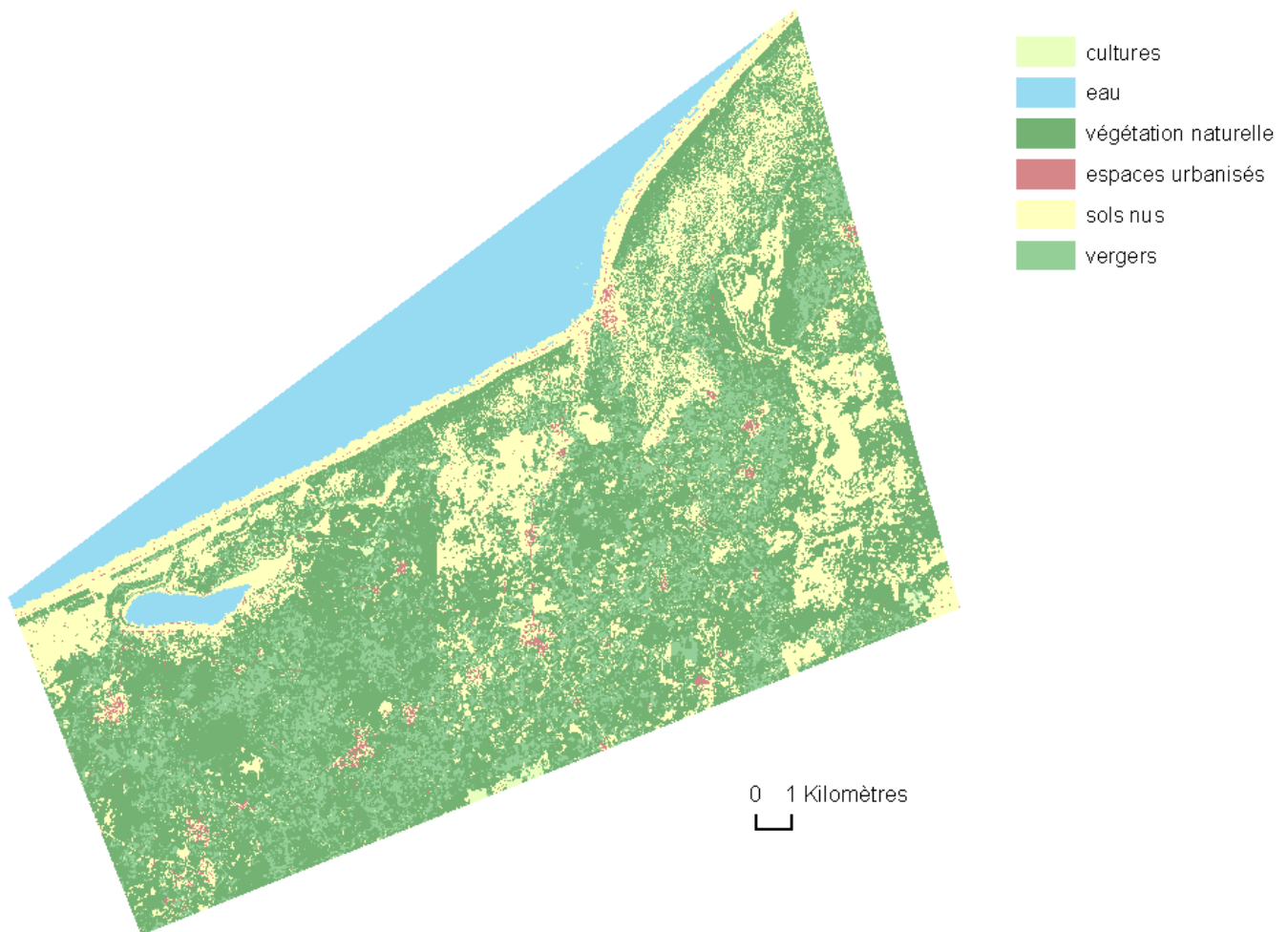


Figure 44 : Classification de l'image totale

Une fois l'image traitée et classifiée, il est temps de se pencher sur les résultats issus de la troisième phase du stage, l'analyse statistique.

C. Effet des pratiques culturelles et du paysage sur le taux d'infestation et le taux de parasitisme

L'analyse des données est le pas décisif vers la compréhension des modèles qui sous-tendent les liens entre le paysage, les pratiques culturelles et l'impact des ravageurs. Ainsi, les résultats seront exploités en suivant le cheminement de la démarche méthodologique, en commençant par établir les connaissances sur les variables, en étudiant les corrélations existant entre les variables, et en étudiant les modèles qui suivent en premier la loi gaussienne et en second la loi binomiale.

1. Connaissances des variables

La connaissance des variables à expliquer et explicatives est importante pour connaître les particularités des variables et ainsi mieux appréhender leur analyse. De manière à limiter le nombre de pages de ce rapport et d'en faciliter la lecture, l'ensemble des histogrammes et des données récapitulatives de chaque variable sont regroupés en annexe 11.

On peut observer que les variables à expliquer suivent globalement des courbes proches de la courbe gaussienne. Néanmoins, des anomalies dans la courbe en cloche des variables tendent à prouver que le lien à la loi normale n'est pas entièrement sûr et nécessite la vérification et dans le cas échéant la mise en place d'une série de lignes de codes d'analyse supplémentaire.

De plus, l'analyse des histogrammes correspondant aux variables culturelles et environnementales ne présentent pas de tracé vraiment identifiables même si certaines variables semblent suivre une répartition linéaire (E7 ou E11 par exemple) ou gaussienne (V1, E5 ou P5). L'analyse des liens de corrélations entre les variables est indispensables afin d'extraire les liens qui existent entre les variables puisqu'ils ne sont pas évidents.

2. Etude des corrélations entre les variables

L'étude des corrélations existantes entre les variables est nécessaire afin de limiter les problèmes liés à une forte corrélation entre deux variables. Il est tout d'abord nécessaire d'établir les liens de corrélation entre les variables à expliquer et ensuite entre ces mêmes variables et les variables explicatives. Finalement, il est cohérent de rechercher des corrélations entre les variables explicatives.

1). Corrélations entre les variables à expliquer

Le premier jeu de corrélations à évaluer est celui entre les variables à expliquer.

La corrélation la plus flagrante est celle entre la variable V1 (Taux de plant infestés) et la variable V4 (nombre de larves par plant) avec un indice de corrélation de 0.84. De plus, la p-value (1-significativité) pour ce couple de variables indique que cette corrélation est grandement significative. Il est donc possible d'éliminer l'une de ces deux variables.

Tableau 10 : Table des corrélations et p-value des variables à expliquer

Corr	V1	V2	V3	V4
V1	1.00	0.10	0.00	0.84
V2	0.10	1.00	-0.08	-0.04
V3	0.00	-0.08	1.00	-0.08
V4	0.84	-0.04	-0.08	1.00

Signif.	V1	V2	V3	V4
V1		0.5993	0.9820	0.9942
V2	0.5993		0.6872	0.8345
V3	0.9820	0.6872		0.6618
V4	0.9942	0.8345	0.6618	

Une fois les variables à expliquer étudiées, il est important de faire également le lien entre ces variables et l'ensemble des variables explicatives, quelles soit reliées aux pratiques culturelles ou à l'environnement. Comme le montre l'annexe 12, aucune corrélation n'apparaît comme significativement importante. Ce qui implique

qu'aucunes des variables explicatives n'est impliqué de manière totalement sûre dans les modèles prédictifs des variables V1 ou V3.

Le lien direct entre les variables explicatives et les variables biologiques à expliquer est peu visible et composé de multiples variables à la fois environnementales et culturelles. Il est néanmoins, important de veiller à l'étude des corrélations pouvant exister entre les variables explicatives.

2). Corrélations entre les variables explicatives

La corrélation entre les différentes variables peut permettre d'éliminer un certain nombre de variables afin d'alléger les analyses et donc de raccourcir les temps de traitements lors de l'analyse. Toujours pour une raison de place et de lisibilité, la table des corrélations présentant les liens entre les variables culturelles et les variables environnementales a été reportée en annexe 13. Les cadres de couleur orange, bleu et vert représentent les différents groupes d'interaction possible (entre les variables culturelles, entre les variables environnementales et entre les variables culturelles et environnementales). Les cellules présentant un fond de couleur verte sont synonyme de forte corrélation

On peut observer qu'aucune corrélation ne ressort comme forte entre une variable culturelle et une variable environnementale.

Les liens entre les variables culturelles sont plus évidents. En effet, trois corrélations sont au-dessus de 0.67 (valeur seuil définie avec les tuteurs). La première fait le lien entre les variables P2 (nombre de traitements phytosanitaires total) et P3 (nombre de traitement phytosanitaires effectuée après le repiquage). Le choix a été fait de conserver la variable P3 car elle correspond aux traitements effectués durant la période d'observation et influence donc directement les données biologiques.

La deuxième corrélation qu'il est important de mettre en évidence est celle entre la variable P3 et la variable P5 (sélectivité). Cette corrélation est à mettre en lien avec celle existant entre les variables P5 et P4 (Type de produits phytosanitaires). Les variables de ce triptyque sont encore une fois mathématiquement corrélées puisque la variable P5 s'exprime par la somme des traitements phytosanitaire pondérée par la sélectivité propre de chaque produit. Ici, c'est donc la variable P4 qui disparaît de l'analyse puisque redondante avec d'autres variables et peu précise (voir annexe 1).

Les liens existants entre les variables environnementales sont plus complexes à étudier. Ceci est dû au fait que ces variables sont plus nombreuses et qu'elles sont imbriquées par groupe de trois (même indicateur à 250, 500 et 1000 m). Ainsi, cinq blocs de corrélations dépassent la valeur seuil de 0.67. La première lie les variables E5 (proportion de végétation naturelle à 250 m) et E6 (proportion de végétation naturelle à 500 m). Le deuxième bloc de corrélations regroupent les variables E10, E11 et E12, correspondant au taux de plantes hôtes dans l'environnement.

Un troisième effet taille influence le bloc composé des variables relatives à l'indice de Shannon (variables E14, E15 et E16). Néanmoins, Le lien de corrélation existant entre les variables E15 et E17 (indice de fragmentation à 250m) est moins évident puisque non expliqué par le phénomène de la taille. L'explication de la relation qui unie ces deux indicateurs est liée à la nature de ces indicateurs et à leur calcul. En effet, l'indice de Shannon prend en compte à la fois la complexité et la fragmentation du milieu. Il est donc plutôt logique de voir cet indice être corrélé avec les indices de fragmentation.

La connaissance des corrélations existantes entre les variables environnementales a permis d'éliminer les variables liées à la fragmentation. Il a été préféré une étude des modèles selon la distance plutôt que l'élimination de variables sur la base des corrélations. Pour cela, deux types de modèles ont été étudié. Le modèle le plus logique à étudier est celui pour lequel les variables suivent une loi normale

3. Recherche des modèles selon la loi normale

Deux variables ont été étudiées dans cette analyse statistique. Il sera donc d'abord étudié les modèles relatifs à la variable V1 (Taux de plat infesté) avant de s'intéresser aux résultats issus de l'étude de la variable V3 (Taux de parasitisme). Pour chacun, il sera discerné les modèles provenant des pratiques culturelles des modèles provenant des variables environnementales avant d'en faire synthèse.

1). Etude du taux d'infestation

L'étude des modèles relatifs aux taux d'infestations se déroule donc en trois étapes, l'étude des pratiques culturelles, celle des variables environnementales et enfin la synthèse de l'ensemble des variables. Pour chaque étude, les trois meilleurs modèles sont présentés.

a. Cultures

Le choix du meilleur modèle se fait selon la valeur du critère informatif d'Akaike (AIC). Ici, c'est le critère dérivé AICc qui sera utilisé puisque l'échantillon est peu important. (Johnson and Omland, 2004). Les trois premiers modèles sont sélectionnés dans chaque cas afin de limiter la perte d'information par le choix d'un modèle sur des critères d'AICc proches.

Tableau 11 : Modèles des pratiques culturelles pour V1

modèle	AICc	p-value		Corr.
P1 P3 P5 P6	-1.19	0.000	***	0.339
P3 P4 P5 P6 P8	-1.68	5.6 ^e -5	***	0.396
P3 P5 P6 P8	-1.93	1.9 ^e -5	***	0.38

Le tableau 11 présente les trois meilleurs modèles pour la prédiction du taux d'infestation à partir des seules variables culturelles. On peut observer que la corrélation entre la prédiction du modèle et les données de l'expérience (colonne 'corr.') est bien en dessous d'un seuil acceptable

même si la p-value indique des valeurs significatives. Il est donc clair que la seule utilisation des variables culturelles ne permet pas de prédire correctement le taux d'infestation.

b. Paysage

Tableau 12 : Modèles des variables du paysage pour V1

N°	Zone	Modèle	AICc	p-value		Corr.
1	250	E10 E14 E20 E4 E7	-0.09	0.004	**	0.34
2	250	E1 E10 E14 E20 E4 E7	0.47	0.005	**	0.23
3	250	E1 E10 E14 E20	-1.06	0.003	**	.032
4	500	E11 E15 E2 E21 E8	-0.36	0.68		0.44
5	500	E11 E15 E2 E21	0.57	0.659		0.319
6	500	E11 E15 E2 E8	2.23	0.006	**	0.377
7	1000	E12 E16 E22 E3 E9	0.28	0.38		0.329
8	1000	E12 E16 E22 E3 E6 E9	0.3	0.448		0.329
9	1000	E16 E22 E3 E6 E9	0.33	0.455		0.323
10	global	E21 E14 E15	-4.1	0.000	***	0.19

Dans le cas de l'étude des modèles issus des variables

environnementales, un problème d'ordre technique et conceptuel s'est posé. En effet, les variables sont regroupés par bloc de trois indicateurs représentant chacun un rayon de recherche différents (250, 500 et 1000m). Après réflexion, il a été

décidé de mené 4 lots de tests sur le paysage correspondant au trois taille de zone tampon (modèles 1 à 9) ainsi qu'à un modèle global (modèle 10) pour les variables environnementales (voir tableau 12). Le modèle globale à été évalué à partir de la matrice restreinte des variables qui étaient ressorties dans les trois meilleurs modèles de chaque zone tampon.

Dans ce tableau, on peut observer que les indices de corrélations des différents modèles prédictifs sont toujours en dessous d'un seul acceptable. Il apparait encore clairement

que les seules variables environnementales sont inaptes à représenter la variabilité du taux d'infestations.

c. Global

Il est possible d'effectuer un dernier jeu de tests à partir des variables issus des meilleurs modèles en ajoutant les variables environnementales aux variables culturelles. A ce stade de l'analyse, seul le meilleur modèle est retenu.

Le modèle qui découle de cette dernière série de lignes de commandes est le suivant :

V1~ 16+E14+E5+E7+P3+P5+P6+P8

Il présente un indice de corrélation de 0.38 avec un p-value non significatif à 0.131.

Ainsi, la recherche de modèle pour le taux d'infestation selon la loi normale n'est pas concluante.

2). Etude du taux de parasitisme

La même étude a été menée pour le taux de parasitisme selon le même protocole

a. Cultures

Le tableau 13 présente les trois meilleurs modèles tirés de l'analyse des variables culturelles en lien avec la variable à

Tableau 13 : Modèles des pratiques culturelles pour V3 (linéaire)

modèle	AICc	p-value		Corr.
P3 P4 P6 P7 P8	-62.8	0.096	.	0.58
P3 P4 P6 P8	-64.9	0.106		0.49
P3 P6 P7 P8	-65.1	0.02	*	0.44

expliquer V3 (Taux de parasitisme). Il apparait que les modèles ont une corrélation plus forte que dans le cas du taux d'infestation. Néanmoins, les p-values sont très peu significatives,

ce qui met un bémol à la valeur de corrélation. Il apparait alors que la seule utilisation des variables culturelles ne permet pas de prédire correctement le taux de parasitisme dans le cas de l'application d'un modèle linéaire normal.

b. Paysage

Tableau 14 : Modèles des variables du paysage pour V3 (linéaire)

num	Modèle	AICc	p-value		Corr.
1	E1 E20 E4	-72.55	0.005	**	0.27
2	E1 E10 E14 E4 E7	-72.86	0.011	**	0.46
3	E1 E10 E14 E20 E4	-73.17	0.005	**	0.51
4	E15 E2 E21 E5 E8	-67.3	0.537		0.45
5	E2 E21 E5 E8	-67.7	0.84		0.48
6	E11 E15 E2 E5 E8	-69	0.15		0.5
7	E12 E16 E22 E3 E9	0.28	0.47		0.577
8	E12 E16 E22 E3 E6 E9	0.3	0.53		0.577
9	E16 E22 E3 E6 E9	0.35	0.50		0.434
10	E8 E9	-64.17	9.1 ^e -9	***	0.08
11	E14 E8	-64.55	6.17 ^e -6	***	0.11
12	E14 E15 E9	64.6	0.01	*	0.21

De même que pour la variable V1, quatre lots d'analyse ont été menés afin de déterminer les modèles pour chaque zone tampon ainsi qu'un modèle global. Dans le tableau 14, on peut observer que les indices de corrélations des différents modèles prédictifs sont moyennement élevés

et donc peu représentatifs. Il apparait encore clairement que les variables environnementales seules sont inaptes à représenter la variabilité du taux de parasitisme.

c. Global

Le dernier lot de tests sur l'ensemble des variables issues des modèles a permis d'extraire le modèle suivant :

$V3 \sim E14 + E15 + E8 + E9 + P3 + P4 + P7 + P5 + P6$
--

Il présente un bon indice de corrélation à 0.67 mais avec un p-value non significatif (0.82). Ainsi, la recherche de modèle pour le taux de parasitisme selon la loi normale n'est pas concluante.

Après discussion avec plusieurs spécialistes de l'analyse statistique du CIRAD, il a été convenu de refaire l'ensemble des textes en appliquant non plus une loi normale mais une loi binomiale.

4. Recherche des modèles selon la loi binomiale

La recherche des modèles prédictifs selon la loi binomiale s'est également effectuée en deux temps, d'abord pour la variable V1 et ensuite pour la variable V3. Pour chaque, le même protocole de tests en trois parties (Pratiques culturelles, paysage et modèles globaux) a été mis en place.

1). Etude du taux d'infestation

Dans le cas de la recherche des modèles suivant la loi binomiale, la variable du taux d'infestation V1 a été transformée en un couple de variables (V11 et V12) représentant le nombre choux est infestés (V11) et le nombre de choux non-infestés (V12) ainsi, la variable V1 se résume par la formule $V1 = V11 / (V11 + V12)$. Avec ses deux nouvelles variables, il est donc possible de mener à terme l'étude des modèles suivant la loi binomiale.

a. Cultures

Tableau 15 : Modèles de pratiques culturelles pour V1 (binomial)

modèle	AICc	p-value		Cor
P3+P7+P8	457,4	8,7 ^e -06	***	0.38
P3+P5+P7+P8	458,4	5,8 ^e -05	***	0.38
P3+P8	458,4	1,9 ^e -05	***	0.36

Le tableau 15 présente les trois meilleurs modèles pour la prédiction du taux d'infestation à partir des seules variables culturelles. On peut observer que la corrélation entre la prédiction du modèle et les données de l'expérience

(colonne 'corr.') est plutôt en dessous d'un seuil acceptable (0.5) même si la p-value indique des valeurs très significatives. Il est donc clair que la seule utilisation des variables culturelles ne permet pas de prédire correctement le taux d'infestation même selon des lois binomiale. Il est donc nécessaire de vérifier si les modèles issus des seules variables environnementales le peuvent.

b. Paysage

Tableau 16 : Modèles de variables de paysages pour V1 (binomial)

Num	Modèle	AICc	p-value		Corr.
1	E1 E14 E4 E7	433.1	7.7 ^e -5	***	0.30
2	E1 E14 E4	435.3	7.5 ^e -5	***	0.30
3	E1 E10 E14 E4 E7	436.0	8.4 ^e -5	***	0.30
4	E15 E2 E21 E5 E8	443.4	0.001	**	0.41
5	E2 E21 E5 E8	445.2	2.8e-5	***	0.41
6	E11 E15 E2 E21 E5 E8	446.8	9.3e-5	***	0.42
7	E16 E22 E9	472.1	7.5e-6	***	0.32
8	E22	472.7	3.9e-6	***	0.29
9	E16 E22	473	3.5e-6	***	0.30
10	E1 E15 E16 E2 E21 E22 E4 E7 E8 E9	340.4	1.21 ^e -15	***	0.60

Dans le cas de l'étude des

modèles issus des variables environnementales, un problème d'ordre technique et conceptuel s'est posé. En effet, les variables sont regroupés par bloc de trois indicateurs représentant chacun un rayon de recherche différents (250, 500 et 1000m). Après

réflexion, il a été décidé de mené 4 lots de tests sur le paysage correspondant au trois taille de zone tampon (modèles 1 à 9) ainsi qu'à un modèle global (modèle 10) pour les variables environnementales. Dans le tableau 16, on peut observer que les indices de corrélations des différents modèles prédictifs sont toujours en dessous d'un seul acceptable sauf pour le dernier modèle. Ceci ci est issus de la matrice contenant l'ensemble des variables ayant été mis en évidence par les trois modèles de chaque taille de zone. Ce modèle particulièrement complexe laisse voir que la recherche des modèles prédictifs par la loi binomiale permet de se rapprocher des données observées sur le terrain. Il est donc temps de voir si la prédiction des modèles issus à la fois des variables culturelles et des variables environnementales permet de réaliser une prédiction encore meilleure.

c. *Global*

Le dernier lot de tests sur l'ensemble des variables issues des modèles a permis d'extraire le modèle suivant :

V1~E15+E21+P3

Il présente un bon indice de corrélation à 0.58 avec un p-value significatif. Ce modèle pourrait donc permettre d'expliquer la variabilité du taux d'infestations en fonction de paramètres à la fois environnementaux comme l'indice de Shannon à 500m (E15) et la complexité à 500m (E21) et cultureux comme le nombre de traitements phytosanitaires après la phase de repiquage (P3).

Tableau 17 : Poids des variables dans le modèle binomial de V1

	OR	2.5 %	97.5 %
Intercept	0.074	0.008	0.663
E15	0.576	0.397	0.833
E21	7.69	2.592	23.077
P3	1.257	1.177	1.342

Le tableau 17 présente les caractéristiques du modèle à travers l'évaluation du Odd Ratio. Cet indicateur permet de visualiser le poids de chaque variable dans le modèle global. Il est ainsi évident que la complexité du paysage (E21) joue un rôle prépondérant dans la définition du modèle, de même que le nombre de traitements sanitaires effectués après le repiquage (P3) et l'indice de

Shannon (E15).

L'apparition de ces deux variables comme déterminantes dans la modélisation s'explique par des faits biologiques. En effet, même si l'individu de *Plutella xylostella* est capable d'effectuer des vols sur de longues distances, il est relativement restreint en termes de plantes hôtes potentiels. Ainsi, un espace complexe lui permet d'espérer trouver des plantes hôtes en plus grandes quantités et donc potentialise un taux d'infestation plus élevé. De plus, les pesticides agissent de manière inverse sur la population de ravageurs en éliminant les individus par vagues et en contrôlant ainsi le taux d'infestation potentiel.

La figure 45 présente ces mêmes résultats sous le jour de l'analyse graphique. Le nuage de point du cinquième graphique montre clairement l'écart existant entre la prédiction (en rouge) et les valeurs observées. De plus, le deuxième graphique tend à démontrer que les données expérimentales concernant le taux d'infestation suivent globalement une loi binomiale (visualisation d'une droite).

Une fois le modèle le plus fiable obtenu pour le taux d'infestation, il est temps d'effectuer la même analyse pour le taux de parasitisme

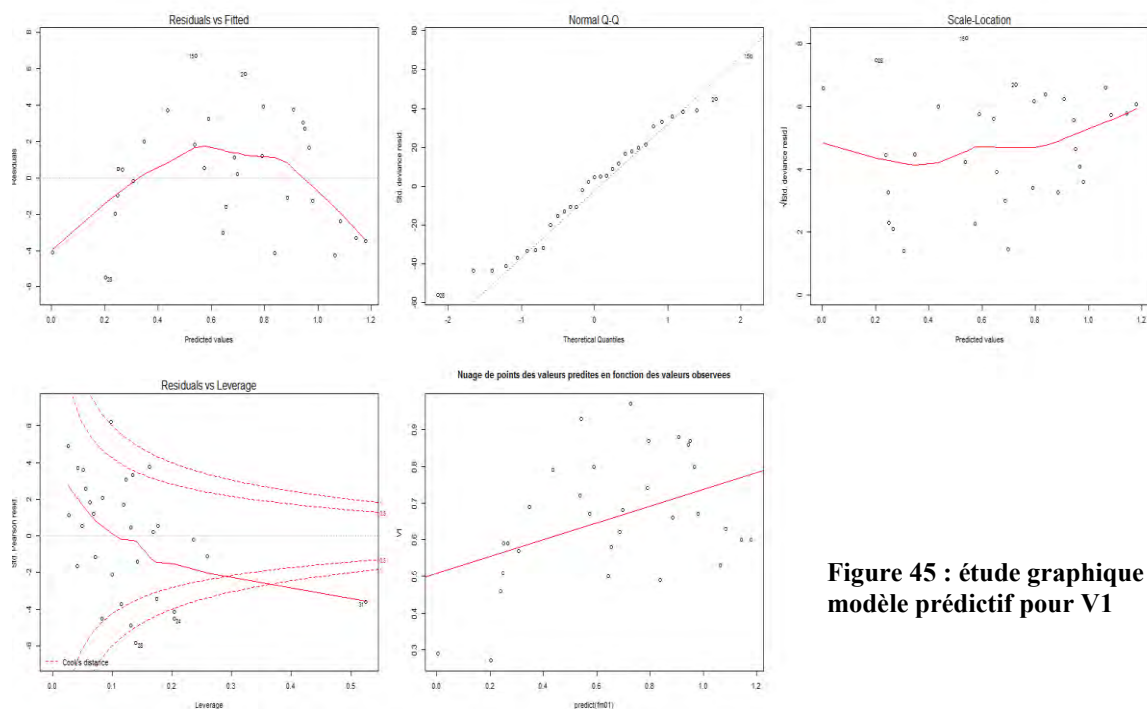


Figure 45 : étude graphique du modèle prédictif pour V1

2). Etude du taux de parasitisme

Dans le cas de la recherche des modèles suivant la loi binomiale, la variable du taux de parasitisme V3 a été transformée en un couple de variables (V31 et V32) représentant le nombre d'individus parasités (V31) et le nombre d'individus non-parasités (V32) ainsi, la variable V3 se résume par la formule $V3 = V31 / (V31 + V32)$. Avec ses deux nouvelles variables, il est donc possible de mener à terme l'étude des modèles suivant la loi binomiale.

a. Cultures

Tableau 18 : Modèles de pratiques culturelles pour V3 (binomial) Le tableau 18 présente les trois meilleurs modèles pour la prédiction du

Modèle	AICc	p-value		Corr.
P3+P5+P6+P7	216,4	2 ^e -16	***	0.49
P1+P3+P5+P7	217,4	2 ^e -16	***	0.47
P3+P5+P7	217,4	2 ^e -16	***	0.46

taux de parasitisme à partir des seules variables culturelles. On peut observer que la corrélation entre la prédiction du modèle et les données de l'expérience est proche de la moitié avec des p-value particulièrement

significatives. Malgré ces deux points positifs, il est donc clair que la seule utilisation des variables culturelles ne permet pas de prédire complètement le taux de parasitisme. Il est donc nécessaire de vérifier si les modèles issus des seules variables environnementales le peuvent.

b. Paysage

Dans le tableau 19, on peut observer que les indices de corrélations des différents modèles prédictifs sont toujours en dessous d'un seul acceptable, même s'ils se rapprochent de la valeur moyenne sauf pour le dernier modèle. Ceci ci est issu de la matrice contenant l'ensemble des variables ayant été mis en évidence par les trois modèles de chaque taille de zone. Ce modèle laisse voir que la recherche des modèles prédictifs par la loi binomiale permet de se rapprocher des données observées sur le terrain. Il est donc temps de voir si la

Tableau 19 : Modèles des indicateurs paysagers pour V3 (binomial)

Num	Modèle	AICc	p-value		Corr.
1	E20+E4	199.1	1.6 ^e -6	***	0.16
2	E14+E20+E4	200.3	7.9e-7	***	0.34
3	E20+E4+E7	201.1	2.2e-6	***	0.46
4	E11+E15+E21	209.5	6.4e-7	***	0.60
5	E11+E21+E5	210	4e-6	***	0.57
6	E11+E21	210.2	4.8e-6	***	0.54
7	E12+E16	208.6	4.6e-11	***	0.57
8	E12+E16+E22	210.8	0.02	*	0.57
9	E12+E16+E6	210.9	4.6e-9	***	0.56
10	E11+E21+E5+E7	197.5	1.4e-6	***	0.67

prédiction des modèles issus à la fois des variables culturelles et des variables environnementales permet de réaliser une prédiction encore meilleure.

c. *Global*

Le dernier lot de tests sur l'ensemble des variables issues des modèles a permis d'extraire le modèle suivant :

$$V3 \sim E11 + E21 + E7 + P5$$

Il présente un très bon indice de corrélation à 0.67 avec un p-value significatif. Ce modèle peut donc permettre d'expliquer la variabilité du taux de parasitisme de manière satisfaisante en fonction de paramètres à la fois environnementaux comme le taux de plantes hôte dans un rayon de 500m (E11), le taux de cultures dans un rayon de 500m (E5), le taux de verger dans un rayon de 250m (E7) et de la complexité à 500m (E21) et culturels comme la note de sélectivité de la parcelle (P5).

Tableau 20 : Poids des variables dans le modèle binomial de V3

	OR	2.5%	97.5 %
Intercept	3.023e-04	1.087e-05	8.478e-03
E11	2.644e+07	2.729e+04	2.438e+10
E21	1.769e+01	3.212e+00	9.527e+01
E5	4.156e+00	1.484e+00	1.152e+01
E7	1.174e+02	5.74e+00	2.368e+03
P5	1.114e+00	9.857e-01	1.259e+00

Le tableau 20 présente les caractéristiques du modèle à travers l'évaluation du Odd Ratio. Il est ici évident que le paramètre le plus important du modèle est le taux de plantes hôtes dans la zone des 500m. Ensuite, ce sont les variables environnementales relatives aux taux de verger (E7) et à la complexité du

paysage (E21) qui pondère le plus le modèle devant l'indice de sélectivité (P5).

La figure 46 présente ces mêmes résultats sous le jour de l'analyse graphique. Le nuage de point du cinquième graphique montre clairement l'écart existant entre la prédiction (en rouge) et les valeurs observées. On peut observer que la dispersion des points est moins importante que dans le cas du taux d'infestation. De plus, le deuxième graphique tend à démontrer que les données expérimentales concernant le taux d'infestation suivent globalement une loi binomiale (visualisation d'une droite).

Une fois ces études effectuées et analysées, il est nécessaire d'apporter un regard critique sur ces résultats et sur les méthodes qui ont permis de les obtenir.

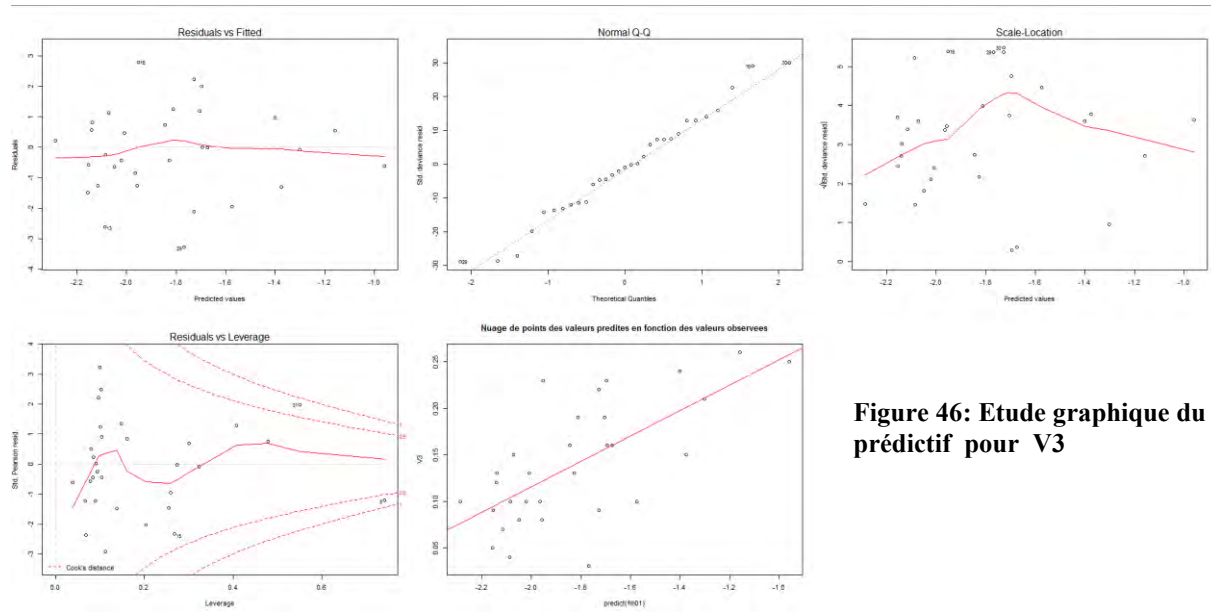


Figure 46: Etude graphique du modèle prédictif pour V3

Discussion

Malgré la grande quantité d'articles et de recherches relatifs à certains des termes du sujet de ce stage, il est apparu lors de l'étude bibliographique que le lien entre la régulation des ravageurs, les pratiques culturales et le paysage est encore loin d'être clairement établi. De plus, le domaine d'étude du ravageur *Plutella xylostella* dans la zone des Niayes au Sénégal n'a encore été que peu exploité. C'est tout d'abord pour cela que cette étude est intéressante. Dans cette partie, il est important d'apporter un regard critique sur les résultats obtenus afin de pouvoir établir les nœuds d'amélioration possible et de dégager les perspectives envisagées après cette étude.

Les premiers résultats obtenus sont plutôt encourageants. En effet, les modèles extraits de l'étude statistiques sont relativement bien corrélés aux variables biologiques. De plus, ils sont composés d'un certains nombres de variables culturales et environnementales expliquées par des phénomènes biologiques déjà identifiés. Ceci a donc permis de confirmer les hypothèses qui sous-tendent cette étude, de caractériser les variables impliquées dans la détermination du taux d'infestation et du taux de parasitisme et de voir que les variables environnementales y ont un rôle prépondérant. Ceci a permis d'éclaircir les connaissances sur les liens existants entre l'écologie, les parasites et les ravageurs. En effet, il a pu être mis en évidence que la forme du paysage avait un impact important sur les ravageurs alors que la composition du paysage était plus importante pour les parasites.

Dans le cas de l'étude du taux d'infestation, trois paramètres sont déterminants. Il s'agit de l'indice de Shannon à 500 m (E15), de la complexité du paysage à 500 m (E21) et du nombre de traitements phytosanitaires après la phase de repiquage (P3). L'apparition de ces trois variables comme déterminantes dans la modélisation s'explique par des faits biologiques. En effet, même si l'individu de *Plutella xylostella* est capable d'effectuer des vols sur de longues distances, il est relativement restreint en termes de plantes hôtes potentiels. Ainsi, un espace complexe lui permet d'espérer trouver des plantes hôtes en plus grandes quantités et donc potentialise un taux d'infestation plus élevé. De plus, les pesticides agissent de manière inverse sur la population de ravageurs en éliminant les individus par vagues et en contrôlant ainsi le taux d'infestation potentiel.

Dans le cas du taux de parasitisme, les paramètres déterminants se rapportent à la composition du paysage avec le taux de plantes hôtes à 500 m (E11), le taux de cultures à 500m (E5), le taux de verger à 250 m (E7) et à sa complexité à 500m (E21) mais aussi aux paramètres culturaux comme la note de sélectivité de la parcelle (P5). L'apparition de ces trois paramètres en tête de la liste des facteurs influençant le taux de parasitisme permet de mieux appréhender les paramètres du milieu important pour les espèces de parasitoïdes. Il apparaît comme évident qu'un paysage riche en arbres (de préférence fruitiers) et en cultures de crucifères tend à favoriser la présence des parasitoïdes et ainsi réguler le taux d'infestation. De plus, un espace complexe semble également augmenter la présence des parasitoïdes en fournissant sans doute un nombre plus important

De plus, la classification mis en place apporte des résultats plutôt satisfaisant. La totalité des classes nécessaires à l'analyse statistiques des données, comme les vergers, la mer ou la végétation naturelle, ont été détectés et identifiés sans trop de confusion. Ceci a permis d'extraire ou de calculer des indicateurs de paysage assez fiables pour être intégré à l'analyse. Enfin, le protocole mis en place sur le terrain et au laboratoire a été bien défini et structurer, ce qui a permis une mise en œuvre rapide avec une méthodologie claire et facilement transposable à d'autres études.

Les résultats obtenus sont encourageants pour la suite puisqu'elle ouvre de nouveaux questionnements. Il sera donc possible de multiplier les observations selon le même protocole

afin d'affiner l'analyse statistique et ainsi d'obtenir des modèles prédictifs plus robustes et plus proches de la réalité sur le terrain. Cette étude a donc pu servir de première base qui sera à étoffer et à compléter pour augmenter la masse de connaissances disponibles sur ce sujet d'étude. De plus, les partenaires rencontrés sur le terrain se sont montrés très intéressés par les résultats de cette étude et par la possibilité de continuer d'aider les chercheurs sénégalais et étrangers dans la compréhension de la régulation des ravageurs.

Même si les résultats de l'étude sont concluants d'un point de vue statistique, un certain nombre d'améliorations sont à envisager afin de rendre ce travail encore plus productif. En effet, les conditions de collecte des données ont été éprouvantes pour l'équipe, d'abord lors de la phase de sélection des parcelles et ensuite lors de la phase de collecte des données servant à l'établissement de la vérité terrain. Ces deux étapes sont plus complexes que compliquées puisque gourmandes en termes de temps de travail et d'énergie déployée. En effet, la phase de récupération de la vérité terrain a duré 15 jours avec un temps de travail proche des 10 heures par jour. Ces contraintes auraient pu être limitées avec la mise en place d'une seconde équipe travaillant en parallèle de la première mais la limitation de cette idée vient du nombre limité de GPS disponibles.

La dépense de temps sur le terrain pour la collecte des données biologiques et culturelles aurait pu être préjudiciable aux autres tâches puisque le temps alloué au stage est limité à 6 mois. Néanmoins, toutes les tâches ont pu être menées à bien même si l'approfondissement de celle-ci n'a pas pu être effectué de manière la plus poussée possible. En effet, la classification affiche des taux de fiabilité faible par rapport au taux acceptable et la confusion entre certaines classes peut être limitée en approfondissant le réglage des valeurs seuils et des paramètres utilisés. De plus, des images de grandes tailles (16 Go chacune) exploitées sur des machines déjà fortement sollicitées ont entraîné des allongements de temps de traitement qui ont ralenti la progression du stage.

C'est d'un point de vue statistique que les résultats obtenus sont les plus probants. En effet, la mise en valeur des variables a permis de révéler les paramètres influençant la biologie de l'insecte ravageur et de ses parasitoïdes et de mieux comprendre les interactions existant entre ces populations et le milieu. Les modèles qui sont issus de cette étude intègrent des paramètres cohérents avec la connaissance empirique existante même si leur robustesse est limitée. Ainsi, le petit nombre de parcelles échantillonnées est une entrave à la significativité des résultats.

Conclusion

L'étude présentée ici a été menée afin de déterminer les liens existant entre les pratiques culturelles, les indicateurs paysagers et la régulation naturelle des insectes ravageurs. Les observations de terrain ont été effectuées dans la zone des Niayes, sur la côte Ouest du Sénégal. Elle porte sur l'espèce *Plutella xylostella* qui est l'un des principaux ravageurs du chou dans cette zone. L'étude a également nécessité une image satellite à très haute résolution spatiale. Les protocoles de collecte des données et de traitement de l'image ont permis de produire des indicateurs environnementaux, biologiques et culturels nécessaires pour la suite de l'étude statistique ainsi qu'un système de classification des images pour cette zone.

Cette étude a permis de valider les hypothèses de départ et de fournir des modèles complets permettant de prédire et de comprendre le taux d'infestation des parcelles observées et le taux de parasitisme des individus collectés. Elle permet donc d'envisager l'ouverture de nouveaux sujets de recherche basée sur les résultats de celle-ci. Néanmoins, cette étude nécessite d'être poursuivie afin d'améliorer à la fois les résultats de la classification mais également les modèles obtenus. Cette étude est donc concluante en termes d'apprentissage qu'ils soient de l'ordre de la méthodologie, de la biologie des espèces ou des perspectives possibles.

Même si elle est encore peu robuste en termes de résultats, cette étude a déjà démontrée son potentiel scientifique et sera sans nul doute réitérée sur la zone des Niayes afin d'engranger de nouvelles connaissances et de poursuivre la recherche de résultats afin d'aider les agriculteurs à produire plus mais surtout à produire mieux.

Bibliographie

Alberg, A.J., Park, J.W., Hager, B.W., Brock, M.V., and Diener-West, M. (2004). The Use of “Overall Accuracy” to Evaluate the Validity of Screening or Diagnostic Tests. *J. Gen. Intern. Med.* 19, 460–465.

Ayalew, G. (2011). Effect of the insect growth regulator novaluron on diamondback moth, *Plutella xylostella* L. (Lepidoptera: Plutellidae), and its indigenous parasitoids. *Crop Prot.* 30, 1087–1090.

Ayalew, G., and Ogol, C.K.P.O. (2006). Occurrence of the diamondback moth (*Plutella xylostella* L.) and its parasitoids in Ethiopia: influence of geographical region and agronomic traits. *J. Appl. Entomol.* 130, 343–348.

Azanza, F., Benjamin, L.R., Brewster, J.L., Craufurd, P.Q., Davis, J.H.C., Drost, D.T., Ellis, R.H., Ewing, E.E., Gray, D., Juvik, J.A., et al. (1997). The physiology of vegetable crops.

Bianchi, F.J.J.A., Goedhart, P.W., and Baveco, J.M. (2008). Enhanced pest control in cabbage crops near forest in The Netherlands. *Landsc. Ecol.* 23, 595–602.

Chaplin-Kramer, R., and Kremen, C. (2012). Pest control experiments show benefits of complexity at landscape and local scales. *Ecol. Appl. Publ. Ecol. Soc. Am.* 22, 1936–1948.

Chaplin-Kramer, R., O’Rourke, M.E., Blitzer, E.J., and Kremen, C. (2011). A meta-analysis of crop pest and natural enemy response to landscape complexity. *Ecol. Lett.* 14, 922–932.

Charleston, D.S., Gols, R., Hordijk, K.A., Kfir, R., Vet, L.E.M., and Dicke, M. (2006). Impact of botanical pesticides derived from *Melia azedarach* and *Azadirachta indica* plants on the emission of volatiles that attract Parasitoids of the diamondback moth to cabbage plants. *J. Chem. Ecol.* 32, 325–349.

Eilers, E.J., and Klein, A.-M. (2009). Landscape context and management effects on an important insect pest and its natural enemies in almond. *Biol. Control* 51, 388–394.

Furlong, M.J., Spafford, H., Ridland, P.M., Endersby, N.M., Edwards, O.R., Baker, G.J., Keller, M.A., and Paull, C.A. (2008). Ecology of diamondback moth in Australian canola: landscape perspectives and the implications for management. *Aust. J. Exp. Agric.* 48, 1494–1505.

Gols, R., Bukovinszky, T., Hemerik, L., Harvey, J.A., Van Lenteren, J.C., and Vet, L.E.M. (2005). Reduced foraging efficiency of a parasitoid under habitat complexity: implications for population stability and species coexistence. *J. Anim. Ecol.* 74, 1059–1068.

Jeger, M.J. (1999). Improved understanding of dispersal in crop pest and disease management: current status and future directions. *Agric. For. Meteorol.* 97, 331–349.

Johnson, J.B., and Omland, K.S. (2004). Model selection in ecology and evolution. *Trends Ecol. Evol.* 19, 101–108.

Landis, R., and Koch, G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 159–174.

Nakahara, L.M., McHugh, J.J., Otsuka, C.K., Funasaki, G.Y., and Lai, P.Y. (1986). Integrated control of diamondback moth and other insect pests using an overhead sprinkler system, an insecticide, and biological control agents, on a watercress farm in Hawaii. (Asian Vegetable Research and Development Center), pp. 403–413.

Qi, J., Kerr, Y., and Chehbouni, A. (1994). External factor consideration in vegetation index development.

Rusch, A., Valantin-Morison, M., Sarthou, J.-P., and Roger-Estrade, J. (2010). Biological Control of Insect Pests in Agroecosystems: Effects of Crop Management, Farming Systems, and Seminatural Habitats at the Landscape Scale: A Review. *Adv. Agron.* 109, 219.

Sarfraz, M., Keddie, A.B., and Dosdall, L.M. (2005). Biological control of the diamondback moth, *Plutella xylostella*: A review Published online 20 July 2005. *Biocontrol Sci. Technol.* 15, 763–789.

Spellerberg, I.F., and Fedor, P.J. (2003). A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the “Shannon–Wiener” Index. *Glob. Ecol. Biogeogr.* 12, 177–179.

Stutz, S., and Entling, M.H. (2011). Effects of the landscape context on aphid-ant-predator interactions on cherry trees. *Biol. Control* 57, 37–43.

Talekar, N.S., and Shelton, A.M. (1993). Biology, Ecology, and Management of the Diamondback Moth. *Annu. Rev. Entomol.* 38, 275–301.

Thies, C., Steffan-Dewenter, I., and Tschardt, T. (2003). Effects of landscape context on herbivory and parasitism at different spatial scales. *Oikos* 101, 18–25.

Veres, A., Petit, S., Conord, C., and Lavigne, C. (2013). Does landscape composition affect pest abundance and their control by natural enemies? A review. *Agric. Ecosyst. Environ.* 166, 110–117.

Wilkinson, T.K., and Landis, D.A. (2005). Habitat diversification in biological control: The role of plant resources. *Plant Provid. Food Carniv. Insects Prot. Mutual. Its Appl. Camb. Univ. Press Camb.* 305–325.